

**STIC-Biotech/ChemLib**

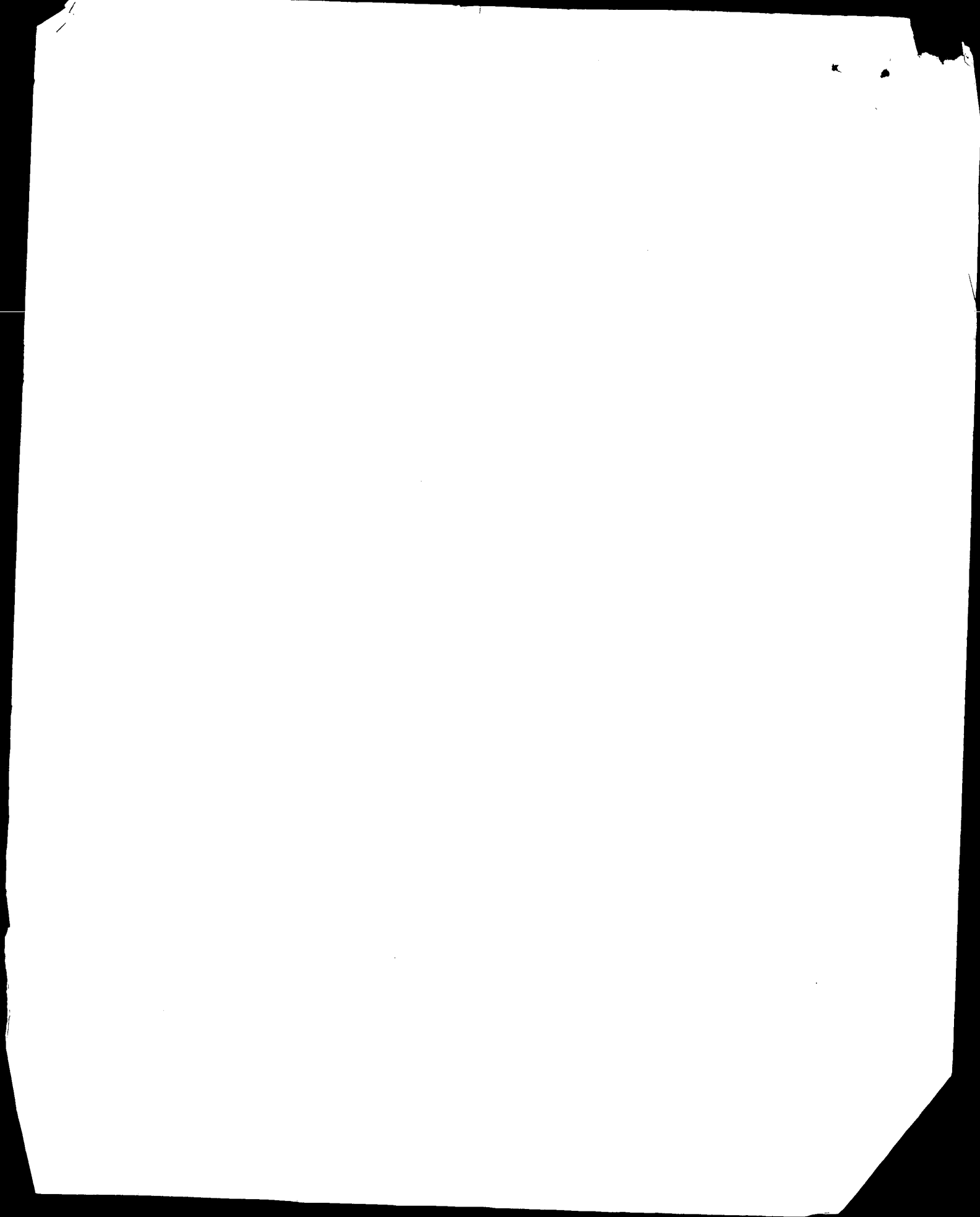
---

**From:** Seharaseyon, Jegatheesan  
**Sent:** Tuesday, July 03, 2001 1:15 PM  
**To:** STIC-Biotech/ChemLib  
**Subject:** Re:09/284320

Please search SEQ ID No: 6 of 09/284,320. Thanks.

J.Seharaseyon  
Art Unit 1647  
CM1 10D16  
10CO1 MB  
(703)-305-1112

09/284,320



**SEARCH REQUEST FORM**

Scientific and Technical Information Center

Requester's Full Name: \_\_\_\_\_ Examiner #: \_\_\_\_\_ Date: \_\_\_\_\_  
 Art Unit: \_\_\_\_\_ Phone Number 30 \_\_\_\_\_ Serial Number: \_\_\_\_\_  
 Mail Box and Bldg/Room Location: \_\_\_\_\_ Results Format Preferred (circle): PAPER DISK E-MAIL

If more than one search is submitted, please prioritize searches in order of need.

\*\*\*\*\*  
 Please provide a detailed statement of the search topic, and describe as specifically as possible the subject matter to be searched. Include the elected species or structures, keywords, synonyms, acronyms, and registry numbers, and combine with the concept or utility of the invention. Define any terms that may have a special meaning. Give examples or relevant citations, authors, etc, if known. Please attach a copy of the cover sheet, pertinent claims, and abstract.

Title of Invention: \_\_\_\_\_

Inventors (please provide full names): \_\_\_\_\_

Earliest Priority Filing Date: \_\_\_\_\_

*\*For Sequence Searches Only\* Please include all pertinent information (parent, child, divisional, or issued patent numbers) along with the appropriate serial number.*

## STAFF USE ONLY

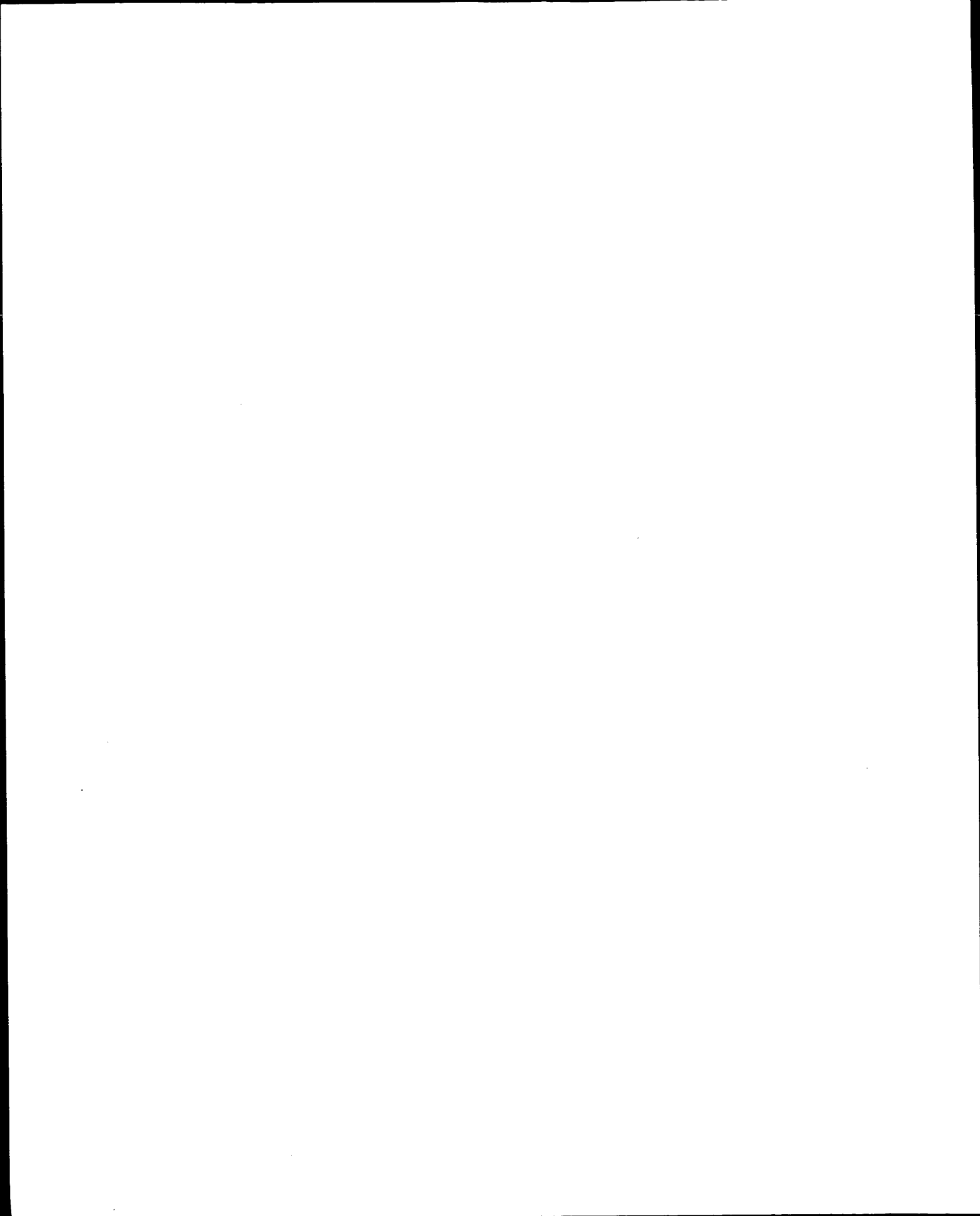
Searcher: Blumenfeld  
 Searcher Phone #: 308-4501  
 Searcher Location: Biotek Lab  
 Date Searcher Picked Up: 7/3/01  
 Date Completed: 7/3/01  
 Searcher Prep & Review Time: \_\_\_\_\_  
 Clerical Prep Time: 3 min  
 Online Time: 2 min

## Type of Search

NA Sequence (#) \_\_\_\_\_  
 AA Sequence (#) 1  
 Structure (#) \_\_\_\_\_  
 Bibliographic \_\_\_\_\_  
 Litigation \_\_\_\_\_  
 Fulltext \_\_\_\_\_  
 Patent Family \_\_\_\_\_  
 Other \_\_\_\_\_

## Vendors and cost where applicable

STN \_\_\_\_\_  
 Dialog \_\_\_\_\_  
 Questel/Orbit \_\_\_\_\_  
 Dr. Link \_\_\_\_\_  
 Lexis/Nexis \_\_\_\_\_  
 Sequence Systems ABSS02  
 WWW/Internet \_\_\_\_\_  
 Other (specify) \_\_\_\_\_



us-09-284-320-6.rag

Tue Jul 3 16:27:27 2001

GenCore version 4.5  
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: July 3, 2001, 16:15:33 ; Search time 22.94 Seconds  
(without alignments)  
924.952 Million cell updates/sec

Title: US-09-284-320-6  
Perfect score: 1763  
Sequence: 1 MAVFVLLALVAGVGLNEFS.....MDPGYDSIIYRMTNQIRMD 350

Scoring table: BLOSUM62  
Gapop 10.0 , Capext 0.5

Searched: 412676 seqs, 60623988 residues  
Total number of hits satisfying chosen parameters: 412676

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000  
Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : A\_Geneseq\_0601.\*  
1: /SID88/gcgdata/geneseq/geneseq/AA1980.DAT.\*  
2: /SID88/gcgdata/geneseq/geneseq/AA1981.DAT.\*  
3: /SID88/gcgdata/geneseq/geneseq/AA1982.DAT.\*  
4: /SID88/gcgdata/geneseq/geneseq/AA1983.DAT.\*  
5: /SID88/gcgdata/geneseq/geneseq/AA1984.DAT.\*  
6: /SID88/gcgdata/geneseq/geneseq/AA1985.DAT.\*  
7: /SID88/gcgdata/geneseq/geneseq/AA1986.DAT.\*  
8: /SID88/gcgdata/geneseq/geneseq/AA1987.DAT.\*  
9: /SID88/gcgdata/geneseq/geneseq/AA1988.DAT.\*  
10: /SID88/gcgdata/geneseq/geneseq/AA1989.DAT.\*  
11: /SID88/gcgdata/geneseq/geneseq/AA1990.DAT.\*  
12: /SID88/gcgdata/geneseq/geneseq/AA1991.DAT.\*  
13: /SID88/gcgdata/geneseq/geneseq/AA1992.DAT.\*  
14: /SID88/gcgdata/geneseq/geneseq/AA1993.DAT.\*  
15: /SID88/gcgdata/geneseq/geneseq/AA1994.DAT.\*  
16: /SID88/gcgdata/geneseq/geneseq/AA1995.DAT.\*  
17: /SID88/gcgdata/geneseq/geneseq/AA1996.DAT.\*  
18: /SID88/gcgdata/geneseq/geneseq/AA1997.DAT.\*  
19: /SID88/gcgdata/geneseq/geneseq/AA1998.DAT.\*  
20: /SID88/gcgdata/geneseq/geneseq/AA1999.DAT.\*  
21: /SID88/gcgdata/geneseq/geneseq/AA2000.DAT.\*  
22: /SID88/gcgdata/geneseq/geneseq/AA2001.DAT.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

# SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1763	100.0	350	19	Human epidermoid c
2	1763	100.0	350	22	Human secreted pro
3	1757	99.7	350	22	Human membrane or
4	1751	99.3	350	22	Human secreted pro
5	1486	84.3	335	20	Secreted protein 3
6	441	25.0	93	20	Human secreted pro
7	121.5	6.9	364	21	Arabidopsis thalia
8	121.5	6.9	392	21	Arabidopsis thalia
9	103.5	5.9	428	19	H. pylori GHPO 137
10	101.5	5.8	318	21	Arabidopsis thalia
11	101	5.7	412	19	H. pylori ORF 09ce

12	100.5	5.7	1103	14	AA39631	Neurofibromatosis
13	100.5	5.7	2485	15	AA39921	RAS associated GAP
14	100.5	5.7	2485	15	AA39922	RAS associated GAP
15	100.5	5.7	2818	13	AA22268	Nfl gene product.
16	100.5	5.7	2818	18	AAW13280	Human neurofibromi
17	97.5	5.5	982	21	AA10105	Feline foamy virus
18	97.5	5.5	982	21	AA10105	Corn leuC subunit
19	96.5	5.5	428	21	AA28433	Rattus norvegicus
20	96	5.4	1886	19	AAW54241	Arabidopsis thalia
21	96	5.4	916	21	AA47980	Arabidopsis thalia
22	96	5.4	922	21	AA47979	Arabidopsis thalia
23	93	5.3	2627	19	AAW61347	Human telomerase R
24	92	5.2	451	21	AA40227	Arabidopsis thalia
25	92	5.2	469	21	AA40226	Arabidopsis thalia
26	92	5.2	517	21	AA40225	Arabidopsis thalia
27	92	5.2	595	19	AAW98283	H. pylori GHPO 125
28	92	5.2	908	20	AA33451	Chimeric Taq DNA p
29	91.5	5.2	454	21	AA53794	Arabidopsis thalia
30	91.5	5.2	511	21	AA53793	Arabidopsis thalia
31	91.5	5.2	545	21	AA53792	Arabidopsis thalia
32	91.5	5.2	789	19	AAW46868	Bacillus thuringie
33	91.5	5.2	789	20	AAV24971	Bacillus thuringie
34	91	5.2	448	21	AA84831	A Streptococcus pn
35	91	5.2	457	21	AA84830	A Streptococcus pn
36	91	5.2	3066	18	AAW44368	Murine Ataxia-tela
37	90.5	5.1	349	19	AAW44368	Aspergillus nidula
38	90.5	5.1	438	21	AAW19277	Arabidopsis thalia
39	90.5	5.1	454	21	AAW19276	Arabidopsis thalia
40	90.5	5.1	454	21	AAW43053	Arabidopsis thalia
41	90.5	5.1	546	21	AAW43052	Arabidopsis thalia
42	90.5	5.1	561	21	AAW43051	Arabidopsis thalia
43	90.5	5.1	611	21	AAW66643	Membrane-bound pro
44	90.5	5.1	673	21	AAW38323	Human secreted pro
45	90.5	5.1	673	21	AAW07428	Amino acid sequenc

## ALIGNMENTS

RESULT 1  
AAW64539  
ID AAW64539 standard; Protein; 350 AA.  
XX  
XX AAW64539;  
AC  
XX  
DT 21-OCT-1998 (first entry)  
XX Human epidermoid carcinoma cell line KB clone HP10034 protein.

XX Transmembrane domain; human; nutrition; cytokine; cell proliferation;  
XX differentiation; immune system; stimulator; suppressor; regulator;  
XX hematopoiesis; activin; inhibitor; chemotactic; chemokine; receptor;  
XX haemostatic; thrombolytic; ligand; anti-inflammatory; tumour.

XX Homo sapiens.  
OS  
XX  
XX WO9821328-A2.  
PN  
XX  
XX 22-MAY-1998.  
PD

XX 07-NOV-1997; 97WO-JP04056.

XX 13-NOV-1996; 96JP-0301429.

XX (PROT-) PROTEGENE INC.  
XX (SAGA ) SAGAMI CHEM RES CENTRE.

XX Kato S, Kobayashi M, Sekine S, Yamaguchi T;

XX WPI; 1998-297932/26.

XX N-PSDB; AA49560, AA49561.

XX Human protein having transmembrane domain - useful for, e.g.

PT

PT research and nutrition  
 XX  
 PS Claim 1; Page 98-99; 205pp; English.  
 XX  
 CC AAM64534-W64558 represent human proteins containing a transmembrane  
 CC domain. These proteins can be used for, e.g. research and nutrition, and  
 CC may have cytokine and cell proliferation/differentiation, immune  
 CC stimulating/suppressing, haematopoiesis regulating, tissue growth,  
 CC activin/inhibin, chemotactic/chemokinetic, haemostatic and thrombolytic,  
 CC receptor/ligand, anti-inflammatory or tumour inhibition activity.  
 XX  
 SQ Sequence 350 AA;

Query Match 100.0%; Score 1763; DB 19; Length 350;  
 Best Local Similarity 100.0%; Pred. No. 8.2e-165;  
 Matches 350; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 1 MAVFVLLALVAGVLGNEFSILKSPGVVFRNGNWPICGERIPDVAAALSMGFVKEDLSW 60  
 DB 1 MAVFVLLALVAGVLGNEFSILKSPGVVFRNGNWPICGERIPDVAAALSMGFVKEDLSW 60  
 QY 61 PGLAVGNLFHRPRATVVMVGVNKLALPPGSSVISYPLENAVPSLDSVANSIHSLSFSEE 120  
 DB 61 PGLAVGNLFHRPRATVVMVGVNKLALPPGSSVISYPLENAVPSLDSVANSIHSLSFSEE 120  
 QY 121 TPVVLQAPSEERYVMGVKANSVFEDLSVTLRQLNRNLFQENSIVLSLPLNSLRNNEVD 180  
 DB 121 TPVVLQAPSEERYVMGVKANSVFEDLSVTLRQLNRNLFQENSIVLSLPLNSLRNNEVD 180  
 QY 181 LFLSELOVLHDISSLLSRHKLAKHSDPLYSLEAGLDEIGKRYGDSQFDRDASKIL 240  
 DB 181 LFLSELOVLHDISSLLSRHKLAKHSDPLYSLEAGLDEIGKRYGDSQFDRDASKIL 240  
 QY 241 VDALKQFADDMYSLYGGNAVVELVTVKSPDTSIRKTRTILEAKQAKNPASPYNLAYKYN 300  
 DB 241 VDALKQFADDMYSLYGGNAVVELVTVKSPDTSIRKTRTILEAKQAKNPASPYNLAYKYN 300  
 QY 301 FEYSVFNVLWIMIALALAVIITSYNIWNMDPGYDSIIYRMTNOKIRMD 350  
 DB 301 FEYSVFNVLWIMIALALAVIITSYNIWNMDPGYDSIIYRMTNOKIRMD 350

RESULT 2  
 AAB90553  
 ID AAB90553 standard; Protein; 350 AA.  
 XX  
 AC AAB90553;  
 XX  
 DT 01-JUN-2001 (first entry)  
 XX  
 DE Human secreted protein, SEQ ID NO: 91.  
 XX  
 KW Human; secreted protein; immunomodulatory; antisclerotic;  
 KW dermatological; anti-inflammatory; anti-HIV; cytostatic; cardiant;  
 KW vascular; anti-angiogenic; ophthalmological; neuroprotectant;  
 KW neutropic; anticonvulsant; antialzheimers; antiparkinsonian;  
 KW antimicrobial; vulnerary; vaccine; gene therapy; cancer;  
 KW protein coordinate data; infection.  
 XX  
 OS Homo sapiens.  
 XX  
 XX WO200121658-A1.  
 PN  
 XX 29-MAR-2001.  
 PD  
 XX 22-SEP-2000; 2000WO-US26013.  
 PF  
 XX 24-SEP-1999; 99US-0155709.  
 PR  
 XX (HUMA-) HUMAN GENOME SCI INC.  
 PA  
 XX Ni J, Baker KP, Birse CE, Ebner R, Fiscella M, Komatsoulis GA;

PI Lafleur DW, Moore PA, Olsen HS, Rosen CA, Ruben SA, Soppet DR;  
 PI Young PE, Wei P, Florence KA;  
 XX  
 DR WPI: 2001-235311/24.  
 XX N-PSDB; AAF97893.  
 PT Nucleic acids encoding 32 human secreted polypeptides, useful for  
 PT preventing, diagnosing and/or treating e.g. cancers, Parkinson's  
 PT disease and diabetic retinopathy -  
 XX  
 PS Claim 11; Page 776-777; 890pp; English.  
 XX

The present sequence is one of 32 novel human secreted polypeptides. The  
 nucleic acid molecules and polypeptides may be used in the prevention,  
 diagnosis and treatment of diseases such as immune disorders (e.g.  
 multiple sclerosis, systemic lupus erythematosus (e.g.  
 immuno-deficiency virus (HIV) infections), hyperproliferative disorders  
 (e.g. cancers and Gaucher's disease), cardiovascular diseases  
 (e.g. Schmitz syndrome, Chaga's cardiomyopathy and coronary  
 arteriosclerosis), angio-genic disorders (e.g. corneal graft  
 neovascularisation), and/or for promoting wound healing, regeneration  
 (e.g. Huntington's chorea, Alzheimer's disease and Parkinson's disease),  
 and/or chemotaxis. The nucleic acid molecules may be used to produce the  
 secreted polypeptides. They may also be used as DNA probes in diagnostic  
 sequences to detect and quantify the presence of similar nucleic acid  
 sequences in samples. The polypeptides may be used as antigens in the  
 production of antibodies and in assays to identify modulators of  
 their expression and activity.

SQ Sequence 350 AA;

Query Match 100.0%; Score 1763; DB 22; Length 350;  
 Best Local Similarity 100.0%; Pred. No. 8.2e-165;  
 Matches 350; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 1 MAVFVLLALVAGVLGNEFSILKSPGVVFRNGNWPICGERIPDVAAALSMGFVKEDLSW 60  
 DB 1 MAVFVLLALVAGVLGNEFSILKSPGVVFRNGNWPICGERIPDVAAALSMGFVKEDLSW 60  
 QY 61 PGLAVGNLFHRPRATVVMVGVNKLALPPGSSVISYPLENAVPSLDSVANSIHSLSFSEE 120  
 DB 61 PGLAVGNLFHRPRATVVMVGVNKLALPPGSSVISYPLENAVPSLDSVANSIHSLSFSEE 120  
 QY 121 TPVVLQAPSEERYVMGVKANSVFEDLSVTLRQLNRNLFQENSIVLSLPLNSLRNNEVD 180  
 DB 121 TPVVLQAPSEERYVMGVKANSVFEDLSVTLRQLNRNLFQENSIVLSLPLNSLRNNEVD 180  
 QY 181 LFLSELOVLHDISSLLSRHKLAKHSDPLYSLEAGLDEIGKRYGDSQFDRDASKIL 240  
 DB 181 LFLSELOVLHDISSLLSRHKLAKHSDPLYSLEAGLDEIGKRYGDSQFDRDASKIL 240  
 QY 241 VDALKQFADDMYSLYGGNAVVELVTVKSPDTSIRKTRTILEAKQAKNPASPYNLAYKYN 300  
 DB 241 VDALKQFADDMYSLYGGNAVVELVTVKSPDTSIRKTRTILEAKQAKNPASPYNLAYKYN 300  
 QY 301 FEYSVFNVLWIMIALALAVIITSYNIWNMDPGYDSIIYRMTNOKIRMD 350  
 DB 301 FEYSVFNVLWIMIALALAVIITSYNIWNMDPGYDSIIYRMTNOKIRMD 350

RESULT 3  
 AAB88347  
 ID AAB88347 standard; Protein; 350 AA.  
 XX  
 AC AAB88347;  
 XX  
 DT 23-MAY-2001 (first entry)  
 XX  
 DE Human membrane or secretory protein clone pSEC0072.  
 XX  
 KW Human; secretory protein; membrane protein; vaccine; gene therapy;

rheumatoid arthritis.  
 Homo sapiens.  
 EPI067182-A2.  
 10-JAN-2001.  
 07-JUL-2000; 2000EP-0114090.  
 08-JUL-1999; 99JP-0194179.  
 11-JAN-2000; 2000JP-0118775.  
 02-MAY-2000; 2000JP-0183766.  
 (HELI-) HELIX RES INST.  
 Ota T., Isogai T, Nishikawa T, Kawai Y, Sugiyama T, Hayashi K;  
 WPI: 2001-053989/11.  
 N-PSDB; AAF93774.  
 Nucleic acids encoding secretory proteins/membrane proteins, useful in  
 gene therapy or as candidate target molecules in drug development -  
 nucleic acids.  
 Claim 1; SEQ ID 62; 609pp + CD ROM; English.  
 This invention relates to nucleic acid sequences AAF93744 - AAF93916  
 which encode human secretory or membrane proteins represented by  
 AAF88317 - AAF88419. Included in the invention are primers to isolate the  
 AAF93917 - AAF94295 and AAF62232. The invention also includes methods for  
 the production of antibodies directed against the proteins and the  
 cDNA sequences, which can be used in vaccines. The polynucleotide sequences  
 can be used in gene therapy. The polynucleotide sequences and the  
 proteins they encode may be used in the prevention, treatment and  
 diagnosis of diseases associated with inappropriate secretory  
 protein/membrane protein expression. The nucleic acids and the  
 proteins may also be used as DNA probes to detect and quantitate the  
 presence of similar nucleic acid sequences in samples. They may also be  
 used as diagnostic agents for detecting the presence of the  
 polypeptides in samples (e.g. by enzyme linked immunosorbent assay  
 (ELISA). Examples of diseases which may be treated include rheumatoid  
 arthritis and diabetes.  
 Sequence 350 AA;  
 Length 350;

[illegible][illegible]

WO200121658-A1.

29-MAR-2001.

22-SEP-2000; 2000WO-US26013.

24-SEP-1999; 99US-0155709.

(HUMA-) HUMAN GENOME SCI INC.

NI J, Baker RP, Birse CE, Ebner R, Fiscella M, Komatsoulis GA; Lafleur DW, Moore PA, Olsen HS, Rosen CA, Ruben SA, Soppet DR; Young PE, Wei P, Florence KA;

WPI: 2001-235311/24.

N-PSDB: AAF97929.

Nucleic acids encoding 32 human secreted polypeptides, useful for preventing, diagnosing and/or treating e.g. cancers, Parkinson's disease and diabetic retinopathy -

Claim 11: Page 809-810: 890pp; English.

The present sequence is one of 32 novel human secreted polypeptides. The nucleic acid molecules and polypeptides may be used in the prevention, diagnosis and treatment of diseases such as immune disorders and multiple sclerosis, systemic lupus erythematosus and human immunodeficiency virus (HIV) infections, hyperproliferative diseases (e.g. cancers and Gaucher's disease), cardiomyopathy and coronary artery disease (e.g. scimitar syndrome, Chaga's cardiomyopathy and coronary artery disease), angiogenic disorders (e.g. corneal graft rejection, arteriosclerosis, and diabetic retinopathy), neurological diseases (e.g. Alzheimer's disease and Parkinson's disease), neurodegeneration (e.g. Huntington's chorea, Alzheimer's disease and Parkinson's disease), neurovascular diseases and/or for promoting wound healing, regeneration (e.g. Huntington's chorea, Alzheimer's disease and Parkinson's disease) and/or chemotaxis. The nucleic acid molecules may be used to produce secreted polypeptides. They may also be used as DNA probes in diagnostic assays to detect and quantify the presence of similar nucleic acid sequences in samples. The polypeptides may be used as antigens in the production of antibodies and in assays to identify modulators of their expression and activity.

Sequence 350 AA;

XX

sequence 350 AA;





```

XX 01-AUG-1997; 97US-0905223.
XX (GEST ) GENSET.
XX Duclert A, Dumas Milne Edwards J, Lacroix B;
XX WPI: 1999-153782/13.
XX DR N-PSDB; AAX51918.
XX
XX New isolated brain-derived nucleic acids - used to develop products
XX PT which may have cytokine, immune, regulatory, haematopoiesis
XX PT regulating, anti-inflammatory or tumour inhibition activity
XX
XX Claim 34; Page 516; 577pp; English.
XX
XX AAX51787 to AAX52019 represent 5' expressed sequence tags (ESTs) for
XX human secreted proteins, and encode the proteins given in AAX12987 to
XX AAX13219, respectively. The proteins given represent the signal peptide
XX and an N-terminal fragment of a secreted protein. The nucleic acid
XX sequences can be used for producing secreted human gene products. They
XX can also be used to develop products for diagnosis and therapy. The
XX proteins obtained may have cytokine activity, cell
XX proliferation/differentiation activity, haematopoiesis regulating
XX activity, tissue growth regulating activity, reproductive hormone
XX regulating activity, chemotactic/chemokinetic activity, haemostatic and
XX thrombolytic activity, receptor/ligand activity, anti-inflammatory
XX activity, tumour inhibition activity or other activities. The products
XX can be used in forensic, gene therapy and chromosome mapping procedures.
XX The sequences can also be used for obtaining corresponding promoter
XX sequences. The nucleic acids encoding the signal peptide can be used for
XX directing extracellular secretion of a polypeptide or the insertion of a
XX polypeptide into a membrane, or importing a polypeptide into a cell.
XX
XX Sequence 93 AA;

Query Match 25.0%; Score 441; DB 20; Length 93;
Best Local Similarity 95.6%; Pred. No. 8.2e-36;
Matches 86; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

QY 1 MAVFVILLALVAGVLGNFSTLKSQGVVFRGWNWPIGRIPDVAALSMGFSVKEDLSW 60
Db 1 MAVFVILLALVAGVLGNFSTLKSQGVVFRGWNWPIGRIPDVAALSMGFSVKEDLSW 60

QY 61 PGLAVGNLFHRPRATVMVMVKGNKLALPP 90
Db 61 PGLAVGNLFHRPRATVMVMVKGNKLALPP 90

RESULT 7
AAG11429
ID AAG11429 standard; Protein; 364 AA.
XX
XX AAG11429;
XX
XX 17-OCT-2000 (first entry)
XX
XX Arabidopsis thaliana protein fragment SEQ ID NO: 10142.
XX
XX Protein identification; signal transduction pathway; metabolic pathway;
XX KW hybridisation assay; genetic mapping; gene expression control; promoter;
XX KW termination sequence.
XX
XX Arabidopsis thaliana.
XX
XX EP1033405-A2.
XX
XX 06-SEP-2000.
XX
XX 25-FEB-2000; 2000EP-0301439.
XX
XX 25-FEB-1999; 99US-0121825.
XX
99US-0123180.
99US-0123548.
99US-0125788.
99US-0126264.
99US-0126785.
99US-0127462.
99US-0128234.
99US-0128714.
99US-0129845.
99US-0130077.
99US-0130449.
99US-0130510.
99US-0130891.
99US-0131449.
99US-0132048.
99US-0132407.
99US-0132484.
99US-0132485.
99US-0132486.
99US-0132487.
99US-0132863.
99US-0134256.
99US-0134218.
99US-0134219.
99US-0134221.
99US-0134370.
99US-0134768.
99US-0134941.
99US-0135124.
99US-0135353.
99US-0135629.
99US-0136021.
99US-0136392.
99US-0136782.
99US-0137222.
99US-0137528.
99US-0137502.
99US-0137724.
99US-0138094.
99US-0138540.
99US-0138847.
99US-0139119.
99US-0139452.
99US-0139453.
99US-0139492.
99US-0139454.
99US-0139455.
99US-0139456.
99US-0139457.
99US-0139458.
99US-0139459.
99US-0139460.
99US-0139461.
99US-0139462.
99US-0139463.
99US-0139750.
99US-0139763.
99US-0139817.
99US-0139899.
99US-0140353.
99US-0140354.
99US-0140695.
99US-0140823.
99US-0140991.
99US-0141287.
99US-0141842.
99US-0142154.
99US-0142055.
99US-0142390.
99US-0142803.
99US-0142920.
99US-0142977.
99US-0143542.

```

PR	14-JUL-1999;	9905-01436224;
PR	15-JUL-1999;	9905-01440051;
PR	16-JUL-1999;	9905-01440085;
PR	17-JUL-1999;	9905-01440086;
PR	18-JUL-1999;	9905-01440325;
PR	19-JUL-1999;	9905-01443311;
PR	20-JUL-1999;	9905-01443332;
PR	21-JUL-1999;	9905-01443333;
PR	22-JUL-1999;	9905-01443334;
PR	23-JUL-1999;	9905-01443335;
PR	24-JUL-1999;	9905-01443522;
PR	25-JUL-1999;	9905-01443523;
PR	26-JUL-1999;	9905-01443524;
PR	27-JUL-1999;	9905-01443526;
PR	28-JUL-1999;	9905-01443518;
PR	29-JUL-1999;	9905-01443519;
PR	30-JUL-1999;	9905-01443515;
PR	31-JUL-1999;	9905-01443521;
PR	01-AUG-1999;	9905-01452218;
PR	02-AUG-1999;	9905-01452128;
PR	03-AUG-1999;	9905-01452128;
PR	04-AUG-1999;	9905-01452704;
PR	05-AUG-1999;	9905-01452913;
PR	06-AUG-1999;	9905-01452918;
PR	07-AUG-1999;	9905-01452918;
PR	08-AUG-1999;	9905-01452918;
PR	09-AUG-1999;	9905-01452918;
PR	10-AUG-1999;	9905-01452918;
PR	11-AUG-1999;	9905-01452918;
PR	12-AUG-1999;	9905-01452918;
PR	13-AUG-1999;	9905-01452918;
PR	14-AUG-1999;	9905-01452918;
PR	15-AUG-1999;	9905-01452918;
PR	16-AUG-1999;	9905-01452918;
PR	17-AUG-1999;	9905-01452918;
PR	18-AUG-1999;	9905-01452918;
PR	19-AUG-1999;	9905-01452918;
PR	20-AUG-1999;	9905-01452918;
PR	21-AUG-1999;	9905-01452918;
PR	22-AUG-1999;	9905-01452918;
PR	23-AUG-1999;	9905-01452918;
PR	24-AUG-1999;	9905-01452918;
PR	25-AUG-1999;	9905-01452918;
PR	26-AUG-1999;	9905-01452918;
PR	27-AUG-1999;	9905-01452918;
PR	28-AUG-1999;	9905-01452918;
PR	29-AUG-1999;	9905-01452918;
PR	30-AUG-1999;	9905-01452918;
PR	31-AUG-1999;	9905-01452918;
PR	01-SEP-1999;	9905-01452918;
PR	02-SEP-1999;	9905-01452918;
PR	03-SEP-1999;	9905-01452918;
PR	04-SEP-1999;	9905-01452918;
PR	05-SEP-1999;	9905-01452918;
PR	06-SEP-1999;	9905-01452918;
PR	07-SEP-1999;	9905-01452918;
PR	08-SEP-1999;	9905-01452918;
PR	09-SEP-1999;	9905-01452918;
PR	10-SEP-1999;	9905-01452918;
PR	11-SEP-1999;	9905-01452918;
PR	12-SEP-1999;	9905-01452918;
PR	13-SEP-1999;	9905-01452918;
PR	14-SEP-1999;	9905-01452918;
PR	15-SEP-1999;	9905-01452918;
PR	16-SEP-1999;	9905-01452918;
PR	17-SEP-1999;	9905-01452918;
PR	18-SEP-1999;	9905-01452918;
PR	19-SEP-1999;	9905-01452918;
PR	20-SEP-1999;	9905-01452918;
PR	21-SEP-1999;	9905-01452918;
PR	22-SEP-1999;	9905-01452918;
PR	23-SEP-1999;	9905-01452918;
PR	24-SEP-1999;	9905-01452918;
PR	25-SEP-1999;	9905-01452918;
PR	26-SEP-1999;	9905-01452918;
PR	27-SEP-1999;	9905-01452918;
PR	28-SEP-1999;	9905-01452918;
PR	29-SEP-1999;	9905-01452918;
PR	30-SEP-1999;	9905-01452918;
PR	01-OCT-1999;	9905-01452918;

PR	05-OCT-1999;	9905-01577553.
PR	06-OCT-1999;	9905-01578665.
PR	07-OCT-1999;	9905-01580829.
PR	08-OCT-1999;	9905-01582332.
PR	12-OCT-1999;	9905-01583639.
PR	13-OCT-1999;	9905-01583923.
PR	13-OCT-1999;	9905-01589294.
PR	13-OCT-1999;	9905-01592595.
PR	14-OCT-1999;	9905-01593232.
PR	14-OCT-1999;	9905-01593330.
PR	14-OCT-1999;	9905-01596337.
PR	14-OCT-1999;	9905-01596384.
PR	18-OCT-1999;	9905-01595884.
PR	21-OCT-1999;	9905-01607411.
PR	21-OCT-1999;	9905-01607657.
PR	21-OCT-1999;	9905-01607658.
PR	21-OCT-1999;	9905-01607770.
PR	21-OCT-1999;	9905-01608115.
PR	21-OCT-1999;	9905-01608415.
PR	22-OCT-1999;	9905-01609801.
PR	22-OCT-1999;	9905-01609811.
PR	22-OCT-1999;	9905-01609899.
PR	25-OCT-1999;	9905-01614004.
PR	25-OCT-1999;	9905-01614005.
PR	25-OCT-1999;	9905-01614006.
PR	26-OCT-1999;	9905-01613539.
PR	26-OCT-1999;	9905-01613661.
PR	26-OCT-1999;	9905-01613661.
PR	28-OCT-1999;	9905-01619220.
PR	28-OCT-1999;	9905-01619922.
PR	28-OCT-1999;	9905-01619923.
PR	29-OCT-1999;	9905-01621442.

Query Match 6.98; Score 121.5; DB 21; Length 364;

Query match	Score	Score	Score	Score
Best Local similarity	22.0%;	Pred. No. 0.0017;		
Matches	86: Conservative	72: Mismatches	154: Indels	79: Gaps
				18:

QY	3	VFWVLLALVAGVGLNFEISILKSPGSVVF-----RNGWNPITGGERIPDPVAALS--	49
		: : : : :     : : : : :     : : : : :     : : : : :	
Db	6	vfvvallltlnyrg-----easgsvffidgsgngylrpssealpmspveisaavaal	59
		: : : : :     : : : : :     : : : : :     : : : : :	
QY	50	MGFSVKREDLSWPG-----LAVGNLPHRPRAVVMVWVKGNKIALPPGSVTSYLENAVPP	103
		: : : : :     : : : : :     : : : : :     : : : : :	
Db	60	lgfapsaatltadgssklunkilpnkpnferpraafvlelagadmlletspshsf_lgnair	118
		: : : : :     : : : : :     : : : : :     : : : : :	
QY	104	PSLDSVANSTHSLFSEETPVVQLALPSEBRYVWCKGNVPEDSVTLFQLNRNL-----	158
		: : : : :     : : : : :     : : : : :     : : : : :	
Db	119	sslks-----dsykadt---elpdnevsvvvnepssdvtckdl--ndfaswlgsgsv	166
		: : : : :     : : : : :     : : : : :     : : : : :	
QY	159	--FOENSJLSSPLNLSNRNEVDL-----LFLSELQVLHDISLSRHKKHLAK--DH	207
		: : : : :     : : : : :     : : : : :     : : : : :	
Db	167	agaepsgllslsiplag-ganvfeulkeaeakfainlligvqirqavsvdyddishgldr	225
		: : : : :     : : : : :     : : : : :     : : : : :	
QY	208	SPDLVSLLEAGLDEIGKRGYSEQDFRASKTILVDALQKADDMWVSLYGGNAVVELVTVK	267
		: : : : :     : : : : :     : : : : :     : : : : :	
Db	226	taeltvyrfggidalaqeyggg--makgmwdvltlisklfnlletshkg---qivgvi	279
		: : : : :     : : : : :     : : : : :     : : : : :	
QY	268	SFDTSLIRKTRILBAQAQKWA-----SPYNLAKYNFESVVFVNNVIM-IA	316
		: : : : :     : : : : :     : : : : :     : : : : :	
Db	280	vldervnqeselnfgssrsarsmvevegipaeaaila-----evilvrltlawltgill	334
		: : : : :     : : : : :     : : : : :     : : : : :	
QY	317	LALAVIITSYNIWNMDPGYSIIYRMTNOKI	347
		: : : : :     : : : : :     : : : : :     : : : : :	
Db	335	lllaatlqgvflmmmlctktdliiv--snvkl	363
		: : : : :     : : : : :     : : : : :     : : : : :	

## RESULT

RESULT  
AAG11428

AA11428  
ID AAG11428 standard: Protein: 392-AA.

XX  
074110WY DT

AC XX

DT 17-OCT-2000 (first entry)  
XX Arabidopsis thaliana protein fragment SEQ ID NO: 10141.  
DE  
XX Protein identification; signal transduction pathway; metabolic pathway;  
KW hybridisation assay; genetic mapping; gene expression control; promoter;  
KW termination sequence.  
XX  
OS Arabidopsis thaliana.  
XX  
PN EP1033405-A2.  
XX  
XX  
PD 06-SEP-2000.  
XX  
XX 25-FEB-2000; 2000EP-0301439.  
PF  
XX 25-FEB-1999; 99US-0121825.  
PR 05-MAR-1999; 99US-0123180.  
PR 09-MAR-1999; 99US-0123348.  
PR 23-MAR-1999; 99US-0125788.  
PR 25-MAR-1999; 99US-0126264.  
PR 29-MAR-1999; 99US-0126785.  
PR 01-APR-1999; 99US-0127462.  
PR 06-APR-1999; 99US-0128234.  
PR 08-APR-1999; 99US-0128714.  
PR 16-APR-1999; 99US-0129845.  
PR 19-APR-1999; 99US-0130077.  
PR 21-APR-1999; 99US-0130449.  
PR 23-APR-1999; 99US-0130510.  
PR 28-APR-1999; 99US-0130891.  
PR 30-APR-1999; 99US-0131449.  
PR 30-APR-1999; 99US-0132048.  
PR 30-APR-1999; 99US-0132407.  
PR 04-MAY-1999; 99US-0132484.  
PR 05-MAY-1999; 99US-0132485.  
PR 06-MAY-1999; 99US-0132486.  
PR 06-MAY-1999; 99US-0132487.  
PR 07-MAY-1999; 99US-0132863.  
PR 11-MAY-1999; 99US-0134256.  
PR 14-MAY-1999; 99US-0134218.  
PR 14-MAY-1999; 99US-0134219.  
PR 14-MAY-1999; 99US-0134321.  
PR 14-MAY-1999; 99US-0134370.  
PR 18-MAY-1999; 99US-0134768.  
PR 19-MAY-1999; 99US-0134941.  
PR 20-MAY-1999; 99US-0135124.  
PR 21-MAY-1999; 99US-0135353.  
PR 24-MAY-1999; 99US-0135629.  
PR 25-MAY-1999; 99US-0136021.  
PR 27-MAY-1999; 99US-0136392.  
PR 28-MAY-1999; 99US-0136782.  
PR 01-JUN-1999; 99US-0137222.  
PR 03-JUN-1999; 99US-0137528.  
PR 04-JUN-1999; 99US-0137502.  
PR 07-JUN-1999; 99US-0137724.  
PR 08-JUN-1999; 99US-0138094.  
PR 10-JUN-1999; 99US-0138540.  
PR 10-JUN-1999; 99US-0138847.  
PR 14-JUN-1999; 99US-0139119.  
PR 16-JUN-1999; 99US-0139452.  
PR 16-JUN-1999; 99US-0139453.  
PR 17-JUN-1999; 99US-0139492.  
PR 18-JUN-1999; 99US-0139454.  
PR 18-JUN-1999; 99US-0139455.  
PR 18-JUN-1999; 99US-0139456.  
PR 18-JUN-1999; 99US-0139457.  
PR 18-JUN-1999; 99US-0139458.  
PR 18-JUN-1999; 99US-0139459.  
PR 18-JUN-1999; 99US-0139460.  
PR 18-JUN-1999; 99US-0139461.  
PR 18-JUN-1999; 99US-0139462.  
PR 18-JUN-1999; 99US-0139463.  
PR 18-JUN-1999; 99US-0139750.  
PR 18-JUN-1999; 99US-0121825.  
PR 18-JUN-1999; 99US-0123180.  
PR 18-JUN-1999; 99US-0123348.  
PR 18-JUN-1999; 99US-0125788.  
PR 18-JUN-1999; 99US-0126264.  
PR 18-JUN-1999; 99US-0126785.  
PR 18-JUN-1999; 99US-0127462.  
PR 18-JUN-1999; 99US-0128234.  
PR 18-JUN-1999; 99US-0128714.  
PR 18-JUN-1999; 99US-0129845.  
PR 18-JUN-1999; 99US-0130077.  
PR 18-JUN-1999; 99US-0130449.  
PR 18-JUN-1999; 99US-0130510.  
PR 18-JUN-1999; 99US-0130891.  
PR 18-JUN-1999; 99US-0131449.  
PR 18-JUN-1999; 99US-0132048.  
PR 18-JUN-1999; 99US-0132407.  
PR 18-JUN-1999; 99US-0132484.  
PR 18-JUN-1999; 99US-0132485.  
PR 18-JUN-1999; 99US-0132486.  
PR 18-JUN-1999; 99US-0132487.  
PR 18-JUN-1999; 99US-0132863.  
PR 18-JUN-1999; 99US-0134256.  
PR 18-JUN-1999; 99US-0134218.  
PR 18-JUN-1999; 99US-0134219.  
PR 18-JUN-1999; 99US-0134321.  
PR 18-JUN-1999; 99US-0134370.  
PR 18-JUN-1999; 99US-0134768.  
PR 18-JUN-1999; 99US-0134941.  
PR 18-JUN-1999; 99US-0135124.  
PR 18-JUN-1999; 99US-0135353.  
PR 18-JUN-1999; 99US-0135629.  
PR 18-JUN-1999; 99US-0136021.  
PR 18-JUN-1999; 99US-0136392.  
PR 18-JUN-1999; 99US-0136782.  
PR 18-JUN-1999; 99US-0137222.  
PR 18-JUN-1999; 99US-0137528.  
PR 18-JUN-1999; 99US-0137502.  
PR 18-JUN-1999; 99US-0137724.  
PR 18-JUN-1999; 99US-0138094.  
PR 18-JUN-1999; 99US-0138540.  
PR 18-JUN-1999; 99US-0138847.  
PR 18-JUN-1999; 99US-0139119.  
PR 18-JUN-1999; 99US-0139452.  
PR 18-JUN-1999; 99US-0139453.  
PR 18-JUN-1999; 99US-0139492.  
PR 18-JUN-1999; 99US-0139454.  
PR 18-JUN-1999; 99US-0139455.  
PR 18-JUN-1999; 99US-0139456.  
PR 18-JUN-1999; 99US-0139457.  
PR 18-JUN-1999; 99US-0139458.  
PR 18-JUN-1999; 99US-0139459.  
PR 18-JUN-1999; 99US-0139460.  
PR 18-JUN-1999; 99US-0139461.  
PR 18-JUN-1999; 99US-0139462.  
PR 18-JUN-1999; 99US-0139463.  
PR 18-JUN-1999; 99US-0139750.  
PR 18-JUN-1999; 99US-0139763.  
PR 18-JUN-1999; 99US-0139817.  
PR 18-JUN-1999; 99US-0139899.  
PR 18-JUN-1999; 99US-0140353.  
PR 18-JUN-1999; 99US-0140354.  
PR 18-JUN-1999; 99US-0140695.  
PR 18-JUN-1999; 99US-0140823.  
PR 18-JUN-1999; 99US-0140991.  
PR 18-JUN-1999; 99US-0141287.  
PR 18-JUN-1999; 99US-0141842.  
PR 18-JUN-1999; 99US-0142154.  
PR 18-JUN-1999; 99US-0142055.  
PR 18-JUN-1999; 99US-0142390.  
PR 18-JUN-1999; 99US-0142803.  
PR 18-JUN-1999; 99US-0142920.  
PR 18-JUN-1999; 99US-0142977.  
PR 18-JUN-1999; 99US-0143342.  
PR 18-JUN-1999; 99US-0143624.  
PR 18-JUN-1999; 99US-0144005.  
PR 18-JUN-1999; 99US-0144085.  
PR 18-JUN-1999; 99US-0144086.  
PR 18-JUN-1999; 99US-0144325.  
PR 18-JUN-1999; 99US-0144331.  
PR 18-JUN-1999; 99US-0144332.  
PR 18-JUN-1999; 99US-0144333.  
PR 18-JUN-1999; 99US-0144334.  
PR 18-JUN-1999; 99US-0144335.  
PR 18-JUN-1999; 99US-0144352.  
PR 18-JUN-1999; 99US-0144632.  
PR 18-JUN-1999; 99US-0144884.  
PR 18-JUN-1999; 99US-0144814.  
PR 18-JUN-1999; 99US-0145086.  
PR 18-JUN-1999; 99US-0145088.  
PR 18-JUN-1999; 99US-0145085.  
PR 18-JUN-1999; 99US-0145087.  
PR 18-JUN-1999; 99US-0145089.  
PR 18-JUN-1999; 99US-0145192.  
PR 18-JUN-1999; 99US-0145145.  
PR 18-JUN-1999; 99US-0145218.  
PR 18-JUN-1999; 99US-0145224.  
PR 18-JUN-1999; 99US-0145276.  
PR 18-JUN-1999; 99US-0145913.  
PR 18-JUN-1999; 99US-0145918.  
PR 18-JUN-1999; 99US-0145919.  
PR 18-JUN-1999; 99US-0145951.  
PR 18-JUN-1999; 99US-0146386.  
PR 18-JUN-1999; 99US-0146388.  
PR 18-JUN-1999; 99US-0146389.  
PR 18-JUN-1999; 99US-0147038.  
PR 18-JUN-1999; 99US-0147204.  
PR 18-JUN-1999; 99US-0147302.  
PR 18-JUN-1999; 99US-0147192.  
PR 18-JUN-1999; 99US-0147260.  
PR 18-JUN-1999; 99US-0147303.  
PR 18-JUN-1999; 99US-0147416.  
PR 18-JUN-1999; 99US-0147493.  
PR 18-JUN-1999; 99US-0147935.  
PR 18-JUN-1999; 99US-0148171.  
PR 18-JUN-1999; 99US-0148319.  
PR 18-JUN-1999; 99US-0148341.  
PR 18-JUN-1999; 99US-0148565.  
PR 18-JUN-1999; 99US-0148684.  
PR 18-JUN-1999; 99US-0149368.  
PR 18-JUN-1999; 99US-0149175.  
PR 18-JUN-1999; 99US-0149426.  
PR 18-JUN-1999; 99US-0149722.  
PR 18-JUN-1999; 99US-0149723.  
PR 18-JUN-1999; 99US-0149929.  
PR 18-JUN-1999; 99US-0149902.  
PR 18-JUN-1999; 99US-0149930.  
PR 18-JUN-1999; 99US-0150566.  
PR 18-JUN-1999; 99US-0150884.  
PR 18-JUN-1999; 99US-0151065.

Qy	3	VVVLLALVAGVGLNERSILKSPGSVVF-----RNGNWPIGPERIDPVAALIS--	49
Db	34	VVVallltlnyr-----easgsvffidsgnnqylrpsaslelmpsrveisaavsal	87
Qy	50	MGFSVKEDLSWPG-----LAVGNLFRPRATVVMVGVKNKIALPPGSVSYPLENAV	103
Db	88	lgtfapsatltagdssklnkikpnqfprpraafvlelagadmlletspshsf-lgnair	146
Qy	104	FSLDSVANSIHLSEETPPVVLQALPSERVYVYGVKANSFEDLSVTLRQLRNRL	158
Db	147	ssiks-----dsykadt---elpdnevvvvsvnepsdvtkdkl--ndfaswlgsyvv	194
Qy	159	--FOENSVLSLPLNSLRNNEVDL-----LFSEQLVHDISSLRKHKIAK--DH	207
Db	195	agaepsgslispiag-ganvefnlekeærkfalnllgylqnirqavsvyddlshgldr	253
Qy	208	SPOLYSLELAGLDEIGRKRYGEQSEDPQASKILYDALQKFADDMYGLYGVNAVVELVTVK	267

22  
23  
24  
25  
26  
27  
28  
29  
30  
31  
32  
33  
34  
35  
36  
37  
38  
39  
40  
41  
42  
43  
44  
45  
46  
47  
48  
49  
50  
51  
52  
53  
54  
55  
56  
57  
58  
59  
60  
61  
62  
63  
64  
65  
66  
67  
68  
69  
70  
71  
72  
73  
74  
75  
76  
77  
78  
79  
80  
81  
82  
83  
84  
85  
86  
87  
88  
89  
90  
91  
92  
93  
94  
95  
96  
97  
98  
99  
100  
101  
102  
103  
104  
105  
106  
107  
108  
109  
110  
111  
112  
113  
114  
115  
116  
117  
118  
119  
120  
121  
122  
123  
124  
125  
126  
127  
128  
129  
130  
131  
132  
133  
134  
135  
136  
137  
138  
139  
140  
141  
142  
143  
144  
145  
146  
147  
148  
149  
150  
151  
152  
153  
154  
155  
156  
157  
158  
159  
160  
161  
162  
163  
164  
165  
166  
167  
168  
169  
170  
171  
172  
173  
174  
175  
176  
177  
178  
179  
180  
181  
182  
183  
184  
185  
186  
187  
188  
189  
190  
191  
192  
193  
194  
195  
196  
197  
198  
199  
200  
201  
202  
203  
204  
205  
206  
207  
208  
209  
210  
211  
212  
213  
214  
215  
216  
217  
218  
219  
220  
221  
222  
223  
224  
225  
226  
227  
228  
229  
230  
231  
232  
233  
234  
235  
236  
237  
238  
239  
240  
241  
242  
243  
244  
245  
246  
247  
248  
249  
250  
251  
252  
253  
254  
255  
256  
257  
258  
259  
260  
261  
262  
263  
264  
265  
266  
267  
268  
269  
270  
271  
272  
273  
274  
275  
276  
277  
278  
279  
280  
281  
282  
283  
284  
285  
286  
287  
288  
289  
290  
291  
292  
293  
294  
295  
296  
297  
298  
299  
300  
301  
302  
303  
304  
305  
306  
307  
308  
309  
310  
311  
312  
313  
314  
315  
316  
317  
318  
319  
320  
321  
322  
323  
324  
325  
326  
327  
328  
329  
330  
331  
332  
333  
334  
335  
336  
337  
338  
339  
340  
341  
342  
343  
344  
345  
346  
347  
348  
349  
350  
351  
352  
353  
354  
355  
356  
357  
358  
359  
360  
361  
362  
363  
364  
365  
366  
367  
368  
369  
370  
371  
372  
373  
374  
375  
376  
377  
378  
379  
380  
381  
382  
383  
384  
385  
386  
387  
388  
389  
390  
391  
392  
393  
394  
395  
396  
397  
398  
399  
400  
401  
402  
403  
404  
405  
406  
407  
408  
409  
410  
411  
412  
413  
414  
415  
416  
417  
418  
419  
420  
421  
422  
423  
424  
425  
426  
427  
428  
429  
430  
431  
432  
433  
434  
435  
436  
437  
438  
439  
440  
441  
442  
443  
444  
445  
446  
447  
448  
449  
450  
451  
452  
453  
454  
455  
456  
457  
458  
459  
460  
461  
462  
463  
464  
465  
466  
467  
468  
469  
470  
471  
472  
473  
474  
475  
476  
477  
478  
479  
480  
481  
482  
483  
484  
485  
486  
487  
488  
489  
490  
491  
492  
493  
494  
495  
496  
497  
498  
499  
500  
501  
502  
503  
504  
505  
506  
507  
508  
509  
510  
511  
512  
513  
514  
515  
516  
517  
518  
519  
520  
521  
522  
523  
524  
525  
526  
527  
528  
529  
530  
531  
532  
533  
534  
535  
536  
537  
538  
539  
540  
541  
542  
543  
544  
545  
546  
547  
548  
549  
550  
551  
552  
553  
554  
555  
556  
557  
558  
559  
560  
561  
562  
563  
564  
565  
566  
567  
568  
569  
570  
571  
572  
573  
574  
575  
576  
577  
578  
579  
580  
581  
582  
583  
584  
585  
586  
587  
588  
589  
590  
591  
592  
593  
594  
595  
596  
597  
598  
599  
600  
601  
602  
603  
604  
605  
606  
607  
608  
609  
610  
611  
612  
613  
614  
615  
616  
617  
618  
619  
620  
621  
622  
623  
624  
625  
626  
627  
628  
629  
630  
631  
632  
633  
634  
635  
636  
637  
638  
639  
640  
641  
642  
643  
644  
645  
646  
647  
648  
649  
650  
651  
652  
653  
654  
655  
656  
657  
658  
659  
660  
661  
662  
663  
664  
665  
666  
667  
668  
669  
670  
671  
672  
673  
674  
675  
676  
677  
678  
679  
680  
681  
682  
683  
684  
685  
686  
687  
688  
689  
690  
691  
692  
693  
694  
695  
696  
697  
698  
699  
700  
701  
702  
703  
704  
705  
706  
707  
708  
709  
710  
711  
712  
713  
714  
715  
716  
717  
718  
719  
720  
721  
722  
723  
724  
725  
726  
727  
728  
729  
730  
731  
732  
733  
734  
735  
736  
737  
738  
739  
740  
741  
742  
743  
744  
745  
746  
747  
748  
749  
750  
751  
752  
753  
754  
755  
756  
757  
758  
759  
760  
761  
762  
763  
764  
765  
766  
767  
768  
769  
770  
771  
772  
773  
774  
775  
776  
777  
778  
779  
780  
781  
782  
783  
784  
785  
786  
787  
788  
789  
790  
791  
792  
793  
794  
795  
796  
797  
798  
799  
800  
801  
802  
803  
804  
805  
806  
807  
808  
809  
810  
811  
812  
813  
814  
815  
816  
817  
818  
819  
820  
821  
822  
823  
824  
825  
826  
827  
828  
829  
830  
831  
832  
833  
834  
835  
836  
837  
838  
839  
840  
841  
842  
843  
844  
845  
846  
847  
848  
849  
850  
851  
852  
853  
854  
855  
85

Db 74 kg lryfemladyerikrmhknkifvdfnfvilggrnigdyfndndtfnfl-dldalff 132

```
QY 191 -----HDIS-SLLSRHKHLAKDHPDLYSLLELACLDIECKRYGEDE 231
Db 133 ggvaskakeserywrfhrspvslrthkrknn-----akeiaklhekipsaedkn 186
QY 232 QFRDASKILVDALQKFADDMYSLYGGNAVVELVTVKSPDTSILIRKTRTILEAKQAKNPAS 291
Db 187 qfekvndfidrfqy---qpylyygnafldadspkldtlyspkikafe-kalknakd 242
QY 292 PYNLAYKY-----NFEYSVFNVMVLMIALALAVIITSYNIWNMDPG 334
Db 243 svfiassfyfpgkmmkfkqngskgieinltn-----sisstdalvvygawe---- 291
QY 335 YDSIYRMNQIRM 349
Db 292 -----tyrnqlvrm 300

RESULT 10
AAGL1430
ID AAGL1430 standard; Protein; 318 AA.
XX
AC AAGL1430;
XX
DT 17-OCT-2000 (first entry)
XX
DE Arabidopsis thaliana protein fragment SEQ ID NO: 10143.
XX
KW Protein identification; signal transduction pathway; metabolic pathway;
KW hybridisation assay; genetic mapping; gene expression control; promoter;
KW termination sequence.
XX
OS Arabidopsis thaliana.
XX
PN EP1033405-A2.
XX
PD 06-SEP-2000.
XX
PF 25-FEB-2000; 2000EP-0301439.
XX
PR 25-FEB-1999; 99US-0121825.
PR 05-MAR-1999; 99US-0123180.
PR 09-MAR-1999; 99US-0123548.
PR 23-MAR-1999; 99US-0125788.
PR 25-MAR-1999; 99US-0126284.
PR 29-MAR-1999; 99US-0126785.
PR 01-APR-1999; 99US-0127462.
PR 06-APR-1999; 99US-0128234.
PR 08-APR-1999; 99US-0128714.
PR 16-APR-1999; 99US-0129845.
PR 19-APR-1999; 99US-0130077.
PR 21-APR-1999; 99US-0130449.
PR 23-APR-1999; 99US-0130510.
PR 23-APR-1999; 99US-0130891.
PR 28-APR-1999; 99US-0131449.
PR 30-APR-1999; 99US-0132048.
PR 30-APR-1999; 99US-0132407.
PR 04-MAY-1999; 99US-0132484.
PR 05-MAY-1999; 99US-0132485.
PR 06-MAY-1999; 99US-0132486.
PR 06-MAY-1999; 99US-0132487.
PR 07-MAY-1999; 99US-0132863.
PR 11-MAY-1999; 99US-0134256.
PR 14-MAY-1999; 99US-0134218.
PR 14-MAY-1999; 99US-0134219.
PR 14-MAY-1999; 99US-0134221.
PR 14-MAY-1999; 99US-0134270.
PR 18-MAY-1999; 99US-0134768.
PR 19-MAY-1999; 99US-0134941.
PR 20-MAY-1999; 99US-0135124.
PR 21-MAY-1999; 99US-0135353.
PR 24-MAY-1999; 99US-0135629.
PR 25-MAY-1999; 99US-0136021.
PR 27-MAY-1999; 99US-0136392.
PR 28-MAY-1999; 99US-0136782.
PR 01-JUN-1999; 99US-0137222.
PR 03-JUN-1999; 99US-0137528.
PR 04-JUN-1999; 99US-0137502.
PR 07-JUN-1999; 99US-0137724.
PR 08-JUN-1999; 99US-0138094.
PR 10-JUN-1999; 99US-0138540.
PR 10-JUN-1999; 99US-0138847.
PR 14-JUN-1999; 99US-0139119.
PR 16-JUN-1999; 99US-0139452.
PR 16-JUN-1999; 99US-0139453.
PR 17-JUN-1999; 99US-0139492.
PR 18-JUN-1999; 99US-0139454.
PR 18-JUN-1999; 99US-0139455.
PR 18-JUN-1999; 99US-0139456.
PR 18-JUN-1999; 99US-0139457.
PR 18-JUN-1999; 99US-0139458.
PR 18-JUN-1999; 99US-0139459.
PR 18-JUN-1999; 99US-0139460.
PR 18-JUN-1999; 99US-0139461.
PR 18-JUN-1999; 99US-0139462.
PR 18-JUN-1999; 99US-0139463.
PR 18-JUN-1999; 99US-0139750.
PR 18-JUN-1999; 99US-0139763.
PR 21-JUN-1999; 99US-0139817.
PR 22-JUN-1999; 99US-0139899.
PR 23-JUN-1999; 99US-0140353.
PR 23-JUN-1999; 99US-0140354.
PR 24-JUN-1999; 99US-0140695.
PR 28-JUN-1999; 99US-0140823.
PR 29-JUN-1999; 99US-0140991.
PR 30-JUN-1999; 99US-0141287.
PR 01-JUL-1999; 99US-0141842.
PR 01-JUL-1999; 99US-0142154.
PR 02-JUL-1999; 99US-0142055.
PR 06-JUL-1999; 99US-0142390.
PR 08-JUL-1999; 99US-0142803.
PR 09-JUL-1999; 99US-0142920.
PR 12-JUL-1999; 99US-0142977.
PR 13-JUL-1999; 99US-0143542.
PR 14-JUL-1999; 99US-0143624.
PR 15-JUL-1999; 99US-0144005.
PR 16-JUL-1999; 99US-0144085.
PR 16-JUL-1999; 99US-0144086.
PR 19-JUL-1999; 99US-0144325.
PR 19-JUL-1999; 99US-0144331.
PR 19-JUL-1999; 99US-0144332.
PR 19-JUL-1999; 99US-0144333.
PR 19-JUL-1999; 99US-0144334.
PR 19-JUL-1999; 99US-0144335.
PR 20-JUL-1999; 99US-0144352.
PR 20-JUL-1999; 99US-0144632.
PR 20-JUL-1999; 99US-0144884.
PR 21-JUL-1999; 99US-0144814.
PR 21-JUL-1999; 99US-0145086.
PR 21-JUL-1999; 99US-0145088.
PR 22-JUL-1999; 99US-0145085.
PR 22-JUL-1999; 99US-0145087.
PR 22-JUL-1999; 99US-0145089.
PR 22-JUL-1999; 99US-0145192.
PR 23-JUL-1999; 99US-0145145.
PR 23-JUL-1999; 99US-0145218.
PR 23-JUL-1999; 99US-0145224.
PR 26-JUL-1999; 99US-0145276.
PR 27-JUL-1999; 99US-0145913.
PR 27-JUL-1999; 99US-0145918.
PR 27-JUL-1999; 99US-0145919.
PR 28-JUL-1999; 99US-0145951.
PR 02-AUG-1999; 99US-0146386.
PR 02-AUG-1999; 99US-0146388.
PR 02-AUG-1999; 99US-0146389.
PR 03-AUG-1999; 99US-0147038.
PR 04-AUG-1999; 99US-0147204.
```

```

PR 04-AUG-1999; 99US-0147302.
PR 05-AUG-1999; 99US-0147192.
PR 05-AUG-1999; 99US-0147260.
PR 06-AUG-1999; 99US-0147303.
PR 06-AUG-1999; 99US-0147416.
PR 09-AUG-1999; 99US-0147493.
PR 09-AUG-1999; 99US-0147935.
PR 10-AUG-1999; 99US-0148171.
PR 11-AUG-1999; 99US-0148319.
PR 12-AUG-1999; 99US-0148341.
PR 13-AUG-1999; 99US-0148565.
PR 13-AUG-1999; 99US-0148684.
PR 16-AUG-1999; 99US-0149368.
PR 17-AUG-1999; 99US-0149175.
PR 18-AUG-1999; 99US-0149426.
PR 20-AUG-1999; 99US-0149722.
PR 20-AUG-1999; 99US-0149723.
PR 20-AUG-1999; 99US-0149929.
PR 23-AUG-1999; 99US-0149902.
PR 23-AUG-1999; 99US-0149920.
PR 25-AUG-1999; 99US-0150366.
PR 26-AUG-1999; 99US-0150884.
PR 27-AUG-1999; 99US-0151065.
PR 27-AUG-1999; 99US-0151066.
PR 27-AUG-1999; 99US-0151080.
PR 30-AUG-1999; 99US-0151303.
PR 31-AUG-1999; 99US-0151438.
PR 01-SEP-1999; 99US-0151930.
PR 07-SEP-1999; 99US-0152363.
PR 10-SEP-1999; 99US-0153070.
PR 13-SEP-1999; 99US-0153758.
PR 15-SEP-1999; 99US-0154018.
PR 16-SEP-1999; 99US-0154039.
PR 20-SEP-1999; 99US-0154779.
PR 22-SEP-1999; 99US-0155139.
PR 23-SEP-1999; 99US-0155486.
PR 24-SEP-1999; 99US-0155659.
PR 28-SEP-1999; 99US-0156458.
PR 29-SEP-1999; 99US-0156596.
PR 04-OCT-1999; 99US-0157117.
PR 05-OCT-1999; 99US-0157753.
PR 06-OCT-1999; 99US-0157865.
PR 07-OCT-1999; 99US-0158029.
PR 08-OCT-1999; 99US-0158232.
PR 12-OCT-1999; 99US-0158369.
PR 13-OCT-1999; 99US-0159293.
PR 13-OCT-1999; 99US-0159294.
PR 13-OCT-1999; 99US-0159295.
PR 14-OCT-1999; 99US-0159329.
PR 14-OCT-1999; 99US-0159330.
PR 14-OCT-1999; 99US-0159331.
PR 14-OCT-1999; 99US-0159637.
PR 14-OCT-1999; 99US-0159638.
PR 18-OCT-1999; 99US-0159584.
PR 21-OCT-1999; 99US-0160741.
PR 21-OCT-1999; 99US-0160767.
PR 21-OCT-1999; 99US-0160768.
PR 21-OCT-1999; 99US-0160770.
PR 21-OCT-1999; 99US-0160814.
PR 21-OCT-1999; 99US-0160815.
PR 22-OCT-1999; 99US-0160980.
PR 22-OCT-1999; 99US-0160981.
PR 22-OCT-1999; 99US-0160989.
PR 25-OCT-1999; 99US-0161404.
PR 25-OCT-1999; 99US-0161405.
PR 25-OCT-1999; 99US-0161406.
PR 26-OCT-1999; 99US-0161359.
PR 26-OCT-1999; 99US-0161360.
PR 26-OCT-1999; 99US-0161361.
PR 28-OCT-1999; 99US-0161920.
PR 28-OCT-1999; 99US-0161922.
PR 28-OCT-1999; 99US-0161993.
PR 29-OCT-1999; 99US-0162142.

Query Match      5.88; Score 101.5; DB 21; Length 318;
Best Local Similarity 21.9%; Pred. No. 0.13;
Matches 75; Conservative 66; Mismatches 141; Indels 61; Gaps 16;

QY 38 PGERIPDVAALSMGFSVKEDLSWPG-----LAVGNLFHRPRATVMVMVKGNKALALPPG 91
   | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 3 pveisaavsal-lgfapsatitadgskinkilknprferpraafvlelagaddmllets 61
   | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
QY 92 SVISYPLENAVPFSLDSVANSIHSLSFSETPVVLQAPSERVYVMGKANSVFEDLSVTL 151
   | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 62 pshsf-lgnairssiks-----dsyked-----elpdnevsvvsvnepsodvtdki-- 108
   | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
QY 152 RQLRNRL-----FQENSVLSSPLNLSLRNNEVDL-----LFLSELQVLHDISLL 197
   | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 109 ndfaswlggsvyagaepsgllsiplag-ganvefnlekeaeerkfainlilgylqniqav 167
   | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
QY 198 SRKHHLAK--DHSPLYSLELAGLDETKRGCESEQFRDASKILYDALOKFADDMYSLY 255
   | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 168 svyodlishgidrtaelvtgrfgidaleqey9gg--makgmdvllistisklfnlletsh 225
   | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
QY 256 GGNVAVELVTKSFDTSILIRKTRTILEAKQAKNPA-----SPYNLAYKYNFEYSV 305
   | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 226 kg-----qivgvilvdervnqesenllnfgssrsarsmvevegipsaalia-----evil 276
   | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
QY 306 VFNMVLWIM-IALALAVIITSYNTWMDPGYDSIIYRWTKNOKI 347
   | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 277 vritawitgilliatilgyvilmumplitkdtily--snvkl 317

RESULT 11
AAY10994
ID AAY10994 standard; Protein; 412 AA.
XX
AC AAY10994;
XX
DT 08-JUN-1999 (first entry)
XX
DE H. pylori ORF 09cel0413_35336707_f2_9 secreted protein.
XX
KW Vaccine; probe; diagnostic; ORF; cell envelope protein;
KW secreted protein; cellular protein.
XX
OS Helicobacter pylori.
XX
PN WO9818323-Al.
XX
PD 07-MAY-1998.
XX
PF 28-OCT-1997; 97WO-US19575.
XX
PR 14-JUL-1997; 97US-0891928.
PR 28-OCT-1996; 96US-0739150.
PR 06-DEC-1996; 96US-0759739.
XX
PA (ASTR ) ASTRA AB.
XX
PI Alm RA, Smith D;
XX
DR WPI; 1998-271811/24.
DR N-PSDB; AAX30461.
XX
PT Helicobacter pylori nucleic acids and proteins - used to develop
PT products for the detection, prevention and treatment of H. pylori
PT infections
XX
PS Claims 27, 31; Page 201-202; 279pp; English.
XX
CC Recombinant or substantially pure preparations of H. pylori polypeptides
CC are disclosed, together with the nucleic acids encoding them. In all,
CC 73 ORFs are shown. The proteins are variously cell envelope proteins,
CC secreted proteins or other cellular proteins. Vaccines containing the
CC nucleic acids or proteins are claimed, as are probes containing at least

```

8 nucleotides from the nucleic acid sequences. The vaccines are useful for treating or reducing the risk of *H. pylori* infections, and the probes can be used diagnostically for detecting the presence of *Helicobacter* in a sample. The products are also of use in screening for compounds having the ability to interfere with the *H. pylori* life cycle or to inhibit *H. pylori* infection.

XX	
SQ	Sequence 412 AA;

Query Match 5.7%; Score 101; DB 19; Length 412;  
Best Local Similarity 20.8%; Pred. No. 0.21;  
Matches 69; Conservative 50; Mismatches 93; Indels 120; Gaps 17;

QY	48	LSWGFVSKEDJUSWGLAVGNLFH-----}	-----RPRATVVMVWVGKNKLAPPGSVI	94
Db	97	mslglsqkvdlngkklctsqxminleqkkllektkqgvinlmingie-----	-----	146
QY	95	SYPLENAVPP-----SLDSVANSIHLSFEETPVWLQAPSE-----ERYVMVGANSVFEDLSV	149	
Db	147	nyknqgeielntaikenlentyqanhabsspdliatleiklleikq-----ndlev	201	
QY	150	TLR-----QURNRLFOENSVLSUPLNSLRNNEVDLLFUSELOVUHDTSLSLSSHKLAK	205	
Db	202	alssshysmgeltfkeneilsiapkn-fefnne-----qelhnisa-----	241	
QY	206	DHSPDLJLSLELAGLDE-----IGRRYGEDSE-----QPRFASKILVDALQKFAADM	251	
Db	242	---tnydiaiarideekaqkditiakksfiedinvtygyfr-----skaynyndm	289	
QY	252	YS-----LYGGNAVVELVTVKSFDT-----	-----SLRKRTFI---	L 281
Db	290	fsvalsiplpylgkqa--kiveqkkeslafksevenakntrhialklkkietiqknl	347	
QY	282	EA-----KOAKNPASPYNLAYKNFEYSVFN	308	
Db	348	esankiikenekiacivaldkknadvnavyn	379	

RESULT	12
AAAR39631	
ID	AAAR39631 standard; Protein; I103 AA.
XX	
AC	AAAR39631;
XX	
DT	16-DEC-1993 (first entry)
XX	
DE	Neurofibromatosis type 1 polypeptide.
XX	
XX	Non-defective gene; NF-1; treatment; tumours; human; detection; ss.

CC The sequence is that of the neurofibromatosis type 1 (NF1)  
 CC polypeptide which may be used therapeutically in the treatment of  
 CC diseases associated with defective NF1 genes, e.g. tumours.  
 XX  
 SQ Sequence 1103 AA;

Query Match	5.7%	Score 100.5;	DB 14;	Length 1103;
Best Local Similarity	22.2%;	Pred. No. 1.1;		
Matches	86;	Conservative 60;	Mismatches 140;	Indels 101; Gaps
<hr/>				
QY	13	GVLGNEFSILKSGSVFVRGNWNPJGE-----RIPDYAALSMGFVSFKEDLSMPCGLAV 65	:	:
DbB	416	gglgskaeavmadtavalsagngkvksvkgmckldktcslpttleqlnmddial 475	:	:
<hr/>				
QY	66	GNLFHRPRATVMVWGKNKLALPPG-----SVTSPLENAVFPISDS-----VANSIHS 115	:	:
DbB	476	-----larymlmsfnssldvaahlpvlfhvvtf-lvatgplsraastghglvinihs 527	:	:
<hr/>				
QY	116	L-----PSEETPWQLAPSE-----ERYVMVGKANSVPFDLSVTIR-QLNRRLFQENSIV 164	:	:
DbB	528	lctcsghlseetkqivrlstefslpkfyllglskvskaaviafrisyrdsfpsgy 587	:	:
<hr/>				
QY	165	-LSSLPLNSLRNNEVDLLFLSELQVLHDISSL-LSRKHHLAK-----DHPDLYSLELA 217	:	:
DbB	588	erctfaltetvea-lleime-acmrdiptckwldqtelaqrfaqynpslqpralv 645	:	:
<hr/>				
QY	218	GLDEIGKR--YGEGSEOFDRASKILVDALQKFADDMYS-----253	:	:
DbB	646	vfgclskrvshgqikqirilskalesclic-gpdytnsqvlieatvialtklqlplinkd 703	:	:
<hr/>				
QY	254	-----LYGGNAVELL-VTVKSFDTSILRKTRITILEAKQAKNPASP-----292	:	:
DbB	704	sphlkalfwavaavlqdvenlysaqtalleqnllhtdisrifndkspeevfmairnple 763	:	:
<hr/>				
QY	293	-----YNLAYKNFEYSVFVNMV 310	:	:
DbB	764	whckqmhdhvafnfnsnfvalvgchll 790	:	:

RESULT	13	
AAAR59921	AAAR59921	standard; protein; 2485 AA.
XX	AAAR59921;	
XX	AC	
XX	AC	
XX	22-FEB-1995	(first entry)
XX	XX	
XX	DE	
XX	XX	
XX	RAS associated GAP NF201.	
XX	Ras: GTPase activating protein; GAP: GAP related domain; GRD;	
XX	PXT10; pKPI1; Saccharomyces cerevisiae; RAS2; v-Ras; heat shock;	
XX	neurofibromatosis type 1; NF1.	
XX	OS	
XX	Homo sapiens.	
XX	WO9416069-A.	
XX	PN	
XX	21-JUL-1994.	
XX	PD	
XX	PPF	12-JAN-1994; 94WO-US00198.
XX	PPR	15-JAN-1993; 93US-0004824.
XX	PA	(SCHE ) SCHERING CORP.
XX	PI	
XX	XX	Kaziro Y, Nakafuku M;
XX	XX	WPI; 1994-249216/30.
XX	DR	
XX	PT	
XX	XX	Blocking Ras-induced effects on a cell - by introducing a GTPase
XX	XX	activating protein to the cell, used esp. in treatment of cancers
XX	XX	

PS Disclosure; Page 36-44; 87pp; English.

XX Human neurofibromatosis type 1 (NF1)-GAP related domain (GRD) was

CC cloned into the yeast expression vector pKTI0 to obtain pKPII. The

CC pKPII DNA was mutagenized by hydroxylamine in vitro and transformed

CC into *S. cerevisiae* TK161-R2V-D, which carries an oncogenic-type

CC RAS2val19 mutation. The heat shock sensitivity of the clones was

CC checked. Plasmid DNAs were recovered, re-transformed into TK161-

CC R2V-D, and phenotypic reversion was examined. 2 Clones, NF201 and

CC NF204 (given in AAR59922), which had strong suppression activity for

CC RAS2val19, were selected. The mutant NF1-GRDs were also able to

CC inhibit v-Ras-induced transformation in mammalian cells.

XX Sequence 2485 AA;

Query Match 5.7%; Score 100.5; DB 15; Length 2485;

Best Local Similarity 22.2%; Pred. No. 3.9;

Matches 86; Conservative 60; Mismatches 140; Indels 101; Gaps 18;

QY 13 GVLGNFESILKSPGVFRNGWNPIGE-----RIPDVAALSMGFSYKEDLSWPGGLAV 65

Db 1667 gglsikaevmadtavalasgnvklsvskvigrmkciidktclstptleghlmwddiai 1726

QY 66 GNLFHRPRATVMVMKGVNKLALPPG-----SVISYPLENAVFPESLDS-----VANSIHS 115

Db 1727 -----larymlmsfnnsldvaahlpylfhvtf-ivatgplslrasthglvinihs 1778

QY 116 L-----FSEETPVVLQAPSE---ERVYMGKANSFEDLSVTLR-QLRNRLFOENSV 164

Db 1779 lctcsqhlseetkqvlrlstefslpkfyllfgiskvksaavlafrssydrsfpsgy 1838

QY 165 -LSSLPLNSLSRNEVDLLFSELQVLHDISL--LSRHKHLAK----DHSPLDLSLELA 217

Db 1839 eretfalsletvtea-lleime-acmrldiptckwldqwtelagrfaqynpslpralv 1896

QY 218 GLDEIGKR--YGESEQFRDASKILVDALQKFADDMYS----- 253

Db 1897 vfgciskrvshgqkqilrlskalesclik--gpdtynsqvlleatvialtklpllnkd 1954

QY 254 -----LYGNAVVEL--VTVKSFDTSLIRKTRTILEAKQKNPASP----- 292

Db 1955 splhkalfwvavavqldevnlysgatalleqnlhtldslrfrndkspveefmairnple 2014

QY 293 -----YNLAYKNFEYSVVFNMV 310

Db 2015 whckqmdhfvgnfnfnfalvgnhll 3041

RESULT 14

AAR59922

ID AAR59922 standard; protein; 2485 AA.

XX AAR59922;

XX 22-FEB-1995 (first entry)

XX RAS associated GAP NF204.

DE

DE

DE

KW Ras: GTPase activating protein; GAP; GAP related domain; GRD;

KW pKTI0; pKPII; *Saccharomyces cerevisiae*; RAS2; v-Ras; heat shock;

KW neurofibromatosis type 1; NF1.

XX Homo sapiens.

XX WO9416069-A.

XX 21-JUL-1994.

XX 12-JAN-1994; 94WO-US00198.

XX 15-JAN-1993; 93US-0004824.

XX

PA (SCHE ) SCHERING CORP.

XX Kaziro Y, Nakafuku M;

XX WPI; 1994-249216/30.

XX Blocking Ras-induced effects on a cell - by introducing a GTPase

PT activating protein to the cell, used esp. in treatment of cancers

PS Disclosure; Page 44-52; 87pp; English.

XX Human neurofibromatosis type 1 (NF1)-GAP related domain (GRD) was

CC cloned into the yeast expression vector pKTI0 to obtain pKPII. The

CC pKPII DNA was mutagenized by hydroxylamine in vitro and transformed

CC into *S. cerevisiae* TK161-R2V-D, which carries an oncogenic-type

CC RAS2val19 mutation. The heat shock sensitivity of the clones was

CC checked. Plasmid DNAs were recovered, re-transformed into TK161-

CC R2V-D, and phenotypic reversion was examined. 2 Clones, NF201

CC (given in AAR59921) and NF204, which had strong suppression activity

CC for RAS2val19, were selected. The mutant NF1-GRDs were also able

CC to inhibit v-Ras-induced transformation in mammalian cells.

XX Sequence 2485 AA;

Query Match 5.7%; Score 100.5; DB 15; Length 2485;

Best Local Similarity 22.2%; Pred. No. 3.9;

Matches 86; Conservative 60; Mismatches 140; Indels 101; Gaps 18;

QY 13 GVLGNFESILKSPGVFRNGWNPIGE-----RIPDVAALSMGFSYKEDLSWPGGLAV 65

Db 1667 gglsikaevmadtavalasgnvklsvskvigrmkciidktclstptleghlmwddiai 1726

QY 66 GNLFHRPRATVMVMKGVNKLALPPG-----SVISYPLENAVFPESLDS-----VANSIHS 115

Db 1727 -----larymlmsfnnsldvaahlpylfhvtf-ivatgplslrasthglvinihs 1778

QY 116 L-----FSEETPVVLQAPSE---ERVYMGKANSFEDLSVTLR-QLRNRLFOENSV 164

Db 1779 lctcsqhlseetkqvlrlstefslpkfyllfgiskvksaavlafrssydrsfpsgy 1838

QY 165 -LSSLPLNSLSRNEVDLLFSELQVLHDISL--LSRHKHLAK----DHSPLDLSLELA 217

Db 1839 eretfalsletvtea-lleime-acmrldiptckwldqwtelagrfaqynpslpralv 1896

QY 218 GLDEIGKR--YGESEQFRDASKILVDALQKFADDMYS----- 253

Db 1897 vfgciskrvshgqkqilrlskalesclik--gpdtynsqvlleatvialtklpllnkd 1954

QY 254 -----LYGNAVVEL--VTVKSFDTSLIRKTRTILEAKQKNPASP----- 292

Db 1955 splhkalfwvavavqldevnlysgatalleqnlhtldslrfrndkspveefmairnple 2014

QY 293 -----YNLAYKNFEYSVVFNMV 310

Db 2015 whckqmdhfvgnfnfnfalvgnhll 2041

RESULT 15

AAR22268

ID AAR22268 standard; Protein; 2818 AA.

XX AAR22268;

XX 06-MAY-1992 (first entry)

XX Nf1 gene product.

XX von Recklinghausen neurofibromatosis disease; autosomal dominant;

XX gene therapy.

XX Homo sapiens.

XX



Search completed: July 3, 2001, 16:18:21  
Job time: 168 sec

7



66 GNT.F

Db 1727 -----LARYMLSLFNNSLDVAHLPYLFHVTF-LVATGPLSLRSTHGLVINIHS 1778  
 Qy 116 L-----FSEETPVVLQAPSE---ERYVMGKANSVFEDLSVTLR-QLNRRLFQENS 164  
 Db 1779 LCTCSQLHSEBTKOVLRLSLTEFSLPKFYLLFGISKVSAAVIAFRSRYDRSFSPGS 1838  
 Qy 165 -LSLPLNSLRNNEVDLLFLSELQVLHDISSL--LSRHKHLAK-----DHSPLDYSLELA 217  
 Db 1839 ERETFALTSLVTEA-LLEIME-ACMRDIPCTCKWLDQWTELAQRAFAQYNPSLOPRALV 1896  
 Qy 218 GLDEIGKR--YGEDESEQROASKILVDLQKFAADMYK-----253  
 Db 1897 VFGCISKRVSHGQIKIIRILSKALESCLK--GPDYNSQVLTAEATVIALTKLOPLLNKD 1954  
 Qy 254 -----LYGGMNAVVEL--VTVKSFDTSLIRKTRTILEAKQAKNPASP-----292  
 Db 1955 SPLHKALFWAVAVLQDLVNLYSAGTALLEQNLHTLDSLRIFNDKSPPEVFAIRNP 2014  
 Qy 293 -----YNLAYKYNFEYSVYFNMV 310  
 Db 2015 WHCKQMDHFVGLNFNSNFALVGHLL 2041

RESULT 4

PCT-US94-00198-2  
 ; Sequence 2, Application PC/TUS9400198  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Schering Corp.  
 ; TITLE OF INVENTION: RAS Associated GAP Proteins  
 ; NUMBER OF SEQUENCES: 6  
 ; CORRESPONDENCE ADDRESS:  
 ; ADDRESSEE: Schering Corp.  
 ; STREET: 1 Giraldo Farms  
 ; CITY: Madison  
 ; STATE: New Jersey  
 ; COUNTRY: USA  
 ; ZIP: 94304-1104  
 ; COMPUTER READABLE FORM:  
 ; MEDIUM TYPE: Floppy disk  
 ; COMPUTER: Macintosh  
 ; OPERATING SYSTEM: 6.0.8  
 ; SOFTWARE: Microsoft Word 5.1a  
 ; CURRENT APPLICATION DATA:  
 ; APPLICATION NUMBER: PCT/US94/00198  
 ; FILING DATE:  
 ; CLASSIFICATION:  
 ; PRIOR APPLICATION DATA:  
 ; APPLICATION NUMBER: US 08/004,824  
 ; FILING DATE: 15-JAN-1993  
 ; ATTORNEY/AGENT INFORMATION:  
 ; NAME: Lunn, Paul G.  
 ; REGISTRATION NUMBER: 32,743  
 ; REFERENCE/DOCKET NUMBER: DX0352 PCT  
 ; TELECOMMUNICATION INFORMATION:  
 ; TELEPHONE: (201)822-7255  
 ; TELEFAX: (201)822-7039  
 ; INFORMATION FOR SEQ ID NO: 2:  
 ; SEQUENCE CHARACTERISTICS:  
 ; LENGTH: 2485 amino acids  
 ; TYPE: amino acid  
 ; STRANDEDNESS: single  
 ; TOPOLOGY: linear  
 ; MOLECULE TYPE: protein  
 ; ORIGINAL SOURCE:  
 ; ORGANISM: Homo sapiens  
 ; PCT-US94-00198-2

Query Match 5.7%; Score 100.5; DB 5; Length 2485;  
 Best Local Similarity 22.28; Pred. No. 1.1;  
 Matches 86; Conservative 60; Mismatches 140; Indels 101; Gaps 18;

Qy 13 GVLGNFSLKSPGVSFVRNGNMPGPE-----RIPDVAALSMGFSVKEDLSWPLAV 65  
 Db 1667 GGLGSIKABVADTAVALASGNVSKVIGRMCKIIDTKTCLSPPTTLEQHLMDIDAI 1726  
 Qy 66 GNLFHRPRATVYVAVKGNKLALPPG-----SVISYPLENAVPSFLDS-----VANSIHS 115  
 Db 1727 -----LARYMLSLFNNSLDVAHLPYLFHVTF-LVATGPLSLRSTHGLVINIHS 1778  
 Qy 116 L-----FSEETPVVLQAPSE---ERYVMGKANSVFEDLSVTLR-QLNRRLFQENS 164  
 Db 1779 LCTCSQLHSEBTKOVLRLSLTEFSLPKFYLLFGISKVSAAVIAFRSRYDRSFSPGS 1838  
 Qy 165 -LSLPLNSLRNNEVDLLFLSELQVLHDISSL--LSRHKHLAK-----DHSPLDYSLELA 217  
 Db 1839 ERETFALTSLVTEA-LLEIME-ACMRDIPCTCKWLDQWTELAQRAFAQYNPSLOPRALV 1896  
 Qy 218 GLDEIGKR--YGEDESEQROASKILVDLQKFAADMYK-----253  
 Db 1897 VFGCISKRVSHGQIKIIRILSKALESCLK--GPDYNSQVLTAEATVIALTKLOPLLNKD 1954  
 Qy 254 -----LYGGMNAVVEL--VTVKSFDTSLIRKTRTILEAKQAKNPASP-----292  
 Db 1955 SPLHKALFWAVAVLQDLVNLYSAGTALLEQNLHTLDSLRIFNDKSPPEVFAIRNP 2014  
 Qy 293 -----YNLAYKYNFEYSVYFNMV 310  
 Db 2015 WHCKQMDHFVGLNFNSNFALVGHLL 2041

RESULT 5

US-08-510-284-1  
 ; Sequence 1, Application US/08510284  
 ; Patent No. 5580955  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Nur-E-kamal, M. S. A.; Maruta, Hiroshi  
 ; TITLE OF INVENTION: FRAGMENTS OF NEUROFIBROMIN (NF1) AND METHOD  
 ; TITLE OF INVENTION: TO REVERSE ACTIVATED RAS INDUCED MALIGNANT TRANSFORMATION I  
 ; TITLE OF INVENTION: MAMMALIAN CELLS  
 ; NUMBER OF SEQUENCES: 2  
 ; CORRESPONDENCE ADDRESS:  
 ; ADDRESSEE: Felfe & Lynch  
 ; STREET: 805 Third Avenue  
 ; CITY: New York City  
 ; STATE: New York  
 ; COUNTRY: USA  
 ; ZIP: 10022  
 ; COMPUTER READABLE FORM:  
 ; MEDIUM TYPE: Diskette, 5.25 inch, 360 kb storage  
 ; COMPUTER: IBM PS/2  
 ; OPERATING SYSTEM: PC-DOS  
 ; SOFTWARE: Wordperfect  
 ; CURRENT APPLICATION DATA:  
 ; APPLICATION NUMBER: US/08/510,284  
 ; FILING DATE:  
 ; CLASSIFICATION: 514  
 ; PRIOR APPLICATION DATA:  
 ; APPLICATION NUMBER: 08/071,575  
 ; FILING DATE: 1-JUNE-1993  
 ; ATTORNEY/AGENT INFORMATION:  
 ; NAME: Hanson, No. 5580955man D.  
 ; REGISTRATION NUMBER: 30,946  
 ; REFERENCE/DOCKET NUMBER: LUD 5319  
 ; TELECOMMUNICATION INFORMATION:  
 ; TELEPHONE: (212) 688-9200  
 ; TELEFAX: (212) 838-3884  
 ; INFORMATION FOR SEQ ID NO: 1:  
 ; SEQUENCE CHARACTERISTICS:  
 ; LENGTH: 2818 amino acids  
 ; TYPE: amino acid  
 ; TOPOLOGY: linear  
 ; MOLECULE TYPE: protein  
 ; PUBLICATION INFORMATION:  
 ; AUTHORS: Marchuk, Douglas A.; Saulino, Ann M.;

; AUTHORS: Tavakkol, Roxanne; Swaroop, Manju;  
 ; AUTHORS: Wallace, Margaret R.; Andersen, Lone B.;  
 ; AUTHORS: Mitchell, Anna L.; Gutmann, David H.;  
 ; AUTHORS: Boguski, Mark; Collins, Francis S.  
 ; TITLE: cDNA Cloning of the Type 1 Neurofibromatosis Gene:  
 ; TITLE: Complete Sequence of the NF1 Gene Product  
 ; Patent No. 5580955  
 ; JOURNAL: Genomics  
 ; VOLUME: 11  
 ; PAGES: 931-940  
 ; DATE: 1991  
 ; US-08-510-284-1

Query Match 5.7%; Score 100.5; DB 1; Length 2818;

Best Local Similarity 22.2%; Pred. No. 1.3;  
Matches 86; Conservative 60; Mismatches 140; Indels 101; Gaps 18;

QY 13 GVLGNEFSILKSPGVVFRNGNWPJGE-----RIPDVAALSMGFSVKEDLSWPGAV 65  
 Db 2000 GGLGSIAEVMADTAVASGNVLYSSKVIKGMCKIIDKTCLSPTTLEQHLWDDIAI 2059  
 QY 66 GNLFRPRATVMVMKGVNKLALPPG-----SVISYPLENAVPSLDS-----VANSIHS 115  
 Db 2060 -----LARYMLMFLSNLDVAHAHLPYLFHVVTFLVATGPLSLRASHGLVNIHS 2111  
 QY 116 L-----FSEETPVVLIQAPSE---ERYVMYKANSVFEDLSVTLR-QLRNLRFQENS 164  
 Db 2112 LCTCSQLHFSEETKQVRLSLTEFSLPKFYLLFGISKVSAAVIAFRSSYRDRSFGSGY 2171  
 QY 165 -LSSLPLNSLSRNEVDLLFLSELQVLHDISL--LSRHKLAK-----DHSPLYSLELA 217  
 Db 2172 EREFTALTSLETVEA-LLEIME-ACWRDIPCKWLDQWTELAQRFQINFSLOPRALV 2229  
 QY 218 GLDEIGKR--YGEDESEQFDASKILVDALQKPADMYS-----253  
 Db 2230 VFGCISKRVSHGQIKQIRILSKALESCLK--GPDYNSQVLIETATVIALTKLOPLLNKD 2287  
 QY 254 -----LYGNAVVEL--VTVKSFDTSLIRKTRTILEAKAKNPASP-----292  
 Db 2288 SPLKALFWAVAVVLQDDEVNLSAGTALLEQNHLTDSLRIFNDKSPPEVFMATRNPLE 2347  
 QY 293 -----YNLAYKYNFEYSVFNVMV 310  
 Db 2348 WHCKQMDHFVGLNFNSNFALVGHLL 2374

## RESULT 6

; US-08-411-389-2  
 ; Sequence 2, Application US/08411389  
 ; Patent No. 5603793

## GENERAL INFORMATION:

; APPLICANT: White, Raymond L.  
 ; APPLICANT: Cawthon, Richard M.  
 ; APPLICANT: Li, Ying  
 ; TITLE OF INVENTION: SOMATIC MUTATIONS IN THE  
 ; TITLE OF INVENTION: NEUROFIBROMATOSIS TYPE 1 GENE IN HUMAN TUMORS  
 ; NUMBER OF SEQUENCES: 21  
 ; CORRESPONDENCE ADDRESS:  
 ; ADDRESSEE: Venable, Baetjer, Howard & Civiletti  
 ; STREET: 1201 New York Avenue NW, Suite 1000  
 ; CITY: Washington  
 ; STATE: DC  
 ; ZIP: 20005  
 ; COMPUTER READABLE FORM:  
 ; MEDIUM TYPE: Floppy disk  
 ; COMPUTER: IBM PC compatible  
 ; OPERATING SYSTEM: PC-DOS/MS-DOS  
 ; SOFTWARE: PatentIn Release #1.0, Version #1.25  
 ; CURRENT APPLICATION DATA:  
 ; APPLICATION NUMBER: US/08/411.389  
 ; FILING DATE:  
 ; CLASSIFICATION: 435

; PRIOR APPLICATION DATA:  
 ; APPLICATION NUMBER: US/08/047,088  
 ; FILING DATE: 16-APR-1993  
 ; ATTORNEY/AGENT INFORMATION:  
 ; NAME: Ihnen, Jeffrey L.  
 ; REGISTRATION NUMBER: 28,957  
 ; REFERENCE/DOCKET NUMBER: 19780-107116  
 ; TELECOMMUNICATION INFORMATION:  
 ; TELEPHONE: 202-962-4810  
 ; TELEFAX: 202-962-8300  
 ; INFORMATION FOR SEQ ID NO: 2:  
 ; SEQUENCE CHARACTERISTICS:  
 ; LENGTH: 2818 amino acids  
 ; TYPE: amino acid  
 ; TOPOLOGY: linear  
 ; MOLECULE TYPE: protein  
 ; US-08-411-389-2

Query Match 5.7%; Score 100.5; DB 1; Length 2818;  
 Best Local Similarity 22.2%; Pred. No. 1.3;  
 Matches 86; Conservative 60; Mismatches 140; Indels 101; Gaps 18;  
 QY 13 GVLGNEFSILKSPGVVFRNGNWPJGE-----RIPDVAALSMGFSVKEDLSWPGAV 65  
 Db 2000 GGLGSIAEVMADTAVASGNVLYSSKVIKGMCKIIDKTCLSPTTLEQHLWDDIAI 2059  
 QY 66 GNLFRPRATVMVMKGVNKLALPPG-----SVISYPLENAVPSLDS-----VANSIHS 115  
 Db 2060 -----LARYMLMFLSNLDVAHAHLPYLFHVVTFLVATGPLSLRASHGLVNIHS 2111  
 QY 116 L-----FSEETPVVLIQAPSE---ERYVMYKANSVFEDLSVTLR-QLRNLRFQENS 164  
 Db 2112 LCTCSQLHFSEETKQVRLSLTEFSLPKFYLLFGISKVSAAVIAFRSSYRDRSFGSGY 2171  
 QY 165 -LSSLPLNSLSRNEVDLLFLSELQVLHDISL--LSRHKLAK-----DHSPLYSLELA 217  
 Db 2172 EREFTALTSLETVEA-LLEIME-ACWRDIPCKWLDQWTELAQRFQINFSLOPRALV 2229  
 QY 218 GLDEIGKR--YGEDESEQFDASKILVDALQKPADMYS-----253  
 Db 2230 VFGCISKRVSHGQIKQIRILSKALESCLK--GPDYNSQVLIETATVIALTKLOPLLNKD 2287  
 QY 254 -----LYGNAVVEL--VTVKSFDTSLIRKTRTILEAKAKNPASP-----292  
 Db 2288 SPLKALFWAVAVVLQDDEVNLSAGTALLEQNHLTDSLRIFNDKSPPEVFMATRNPLE 2347  
 QY 293 -----YNLAYKYNFEYSVFNVMV 310  
 Db 2348 WHCKQMDHFVGLNFNSNFALVGHLL 2374

## RESULT 7

; US-08-449-933-2  
 ; Sequence 2, Application US/08449933  
 ; Patent No. 5859195

## GENERAL INFORMATION:

; APPLICANT: Collins, Francis S.  
 ; APPLICANT: Wallace, Margaret R.  
 ; APPLICANT: Marchuk, Douglas A.  
 ; APPLICANT: Anderson, Lone B.  
 ; APPLICANT: Guttman, David H.  
 ; TITLE OF INVENTION: Neurofibromatosis Gene  
 ; NUMBER OF SEQUENCES: 14  
 ; CORRESPONDENCE ADDRESS:  
 ; ADDRESSEE: Morrison & Foerster  
 ; STREET: 755 Page Mill Road  
 ; CITY: Palo Alto  
 ; STATE: CA  
 ; COUNTRY: USA  
 ; ZIP: 94304-1018  
 ; COMPUTER READABLE FORM: disk  
 ; MEDIUM TYPE: Floppy disk

```

; OTHER INFORMATION: acid insertion(SEQ ID NO:10) representing an alternatively
; FEATURE:
; NAME/KEY: Modified-site
; LOCATION: (1370*1371)
; OTHER INFORMATION: /note= "Position of a 21 amino acid
; OTHER INFORMATION: Insertion representing an alternatively spliced product"
; FEATURE:
; NAME/KEY: Domain
; LOCATION: 1125...1137
; OTHER INFORMATION: /note= "Nfl catalytic domain"
; FEATURE:
; NAME/KEY: Modified-site
; LOCATION: 2746...2818
; OTHER INFORMATION: /note= "Corresponding amino acids
; OTHER INFORMATION: for the PstI-HindIII fragment designated pMAL.B3A"
; FEATURE:
; NAME/KEY: Modified-site
; LOCATION: 65..1240
; OTHER INFORMATION: /note= "Corresponding amino acids
; OTHER INFORMATION: for the HpaI-XhoI fragment designated pMAL.HF3A.X"
; PUBLICATION INFORMATION:
; AUTHORS: Wallace, M.R. et al.
; TITLE: Type 1 Neurofibromatosis Gene: Correction
; JOURNAL: Science
; VOLUME: 250
; ISSUE: 12/21/90
; PAGES: 1749-
; DATE: 12/21-1990
; RELEVANT RESIDUES IN SEQ ID NO: 2: FROM 1 TO 2818
; PUBLICATION INFORMATION:
; AUTHORS: Wallace, M.R. et al.
; TITLE: Type 1 Neurofibromatosis Gene: Identification
; TITLE: of a Large Transcript in Three NFL Patients
; JOURNAL: Science
; VOLUME: 249
; ISSUE: 07/13/90
; PAGES: 181-186
; DATE: 07/13-1990
; RELEVANT RESIDUES IN SEQ ID NO: 2: FROM 1 TO 2818
; US-08-449-933-2

Query Match 5.7%; Score 100.5; DB 2; Length 2818;
Best Local Similarity 22.2%; Pred. No. 1.3;
Matches 86; Conservative 60; Mismatches 140; Indels 101; Gaps 18;

QY 13 GVIGNEFSLKSGVVRNGNWPTPG-----RIPDVAALSGFSVKEDLSWPGGLAV 65
Db 2000 GGLGTSKAEWADTAVALASGNVKSIVGRMKCIKDICTLPTTLEQHLMMDDIAI 2059
QY 66 GNLFHRPRATVWVGVKNKLPPG-----SVISYPLENAVPFSLDS-----VANSHRS 115
Db 2060 -----LARYMLMSFNNSLDVAHLPLFHVVTFLVATGPLSLRASHGLVINIHS 2111
QY 116 L-----FSEETPVQLQAPSE-----ERVWVGKANSFEDLSVTLR-QLRNRLQENSV 164
Db 2112 LCTCSLHFSEETQVLRSLTEFLPKPFLYLLFGISKVSAVIAFRSSYRDRSPSGY 2171
QY 165 -LSLPLNSLSRNEEDLLFLSELQVLDHDISSL--LSRHKHLAK-----DHSPLYSLELA 217
Db 2172 ERETFTALSLTEVTEA-LLEIME-ACMEDIPTCKWLDQWTELAQFAQFNLSLPALV 2229
QY 218 GLDEIGKR--YGEDSEQFDKASKILVDALQKFADDMYS----- 253
Db 2230 VFGICISKRVSHGQIKIIRLLSKALESCLK--GPDYNSQVLIETVIALTKLOPLLNKD 2287
QY 254 -----LYGGNAVVEL--VTVKSFDTSLIRKRTILEAKOAKNPSP-----292

```

Db 2288 SPLKALFWAVLQDEVNLYSAGTALLEQNLTLDSLRIFNDKSPFEEVFWAIRNPLE 2347  
QY 293 -----YNLAKYNFEYSVFNWV 310  
Db 2348 WHCKQMDHFVGLNFNSNFALVGHLL 2374

## RESULT

US-07-966-049A-2  
Sequence 2, Application US/07966049A  
Patent No. 6238861  
GENERAL INFORMATION:  
APPLICANT: Collins, Francis S.  
APPLICANT: Wallace, Margaret R.  
APPLICANT: Marchuk, Douglas A.  
APPLICANT: Anderson, Lone B.  
APPLICANT: Guttman, David H.  
TITLE OF INVENTION: Neurofibromatosis Gene  
NUMBER OF SEQUENCES: 14  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Morrison & Foerster  
STREET: 755 Page Mill Road  
CITY: Palo Alto  
STATE: CA  
COUNTRY: USA  
ZIP: 94304-1018  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/07/966,049A  
FILING DATE: 02-AUG-1993  
CLASSIFICATION: 435  
ATTORNEY/AGENT INFORMATION:  
NAME: Kanski, Antoinette F.  
REGISTRATION NUMBER: 34,202  
REFERENCE/DOCKET NUMBER: 20344-20553.20  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (415) 813-5600  
TELEFAX: (415) 494-0792  
TELEX: 706141 MRSNFOERS SFO  
INFORMATION FOR SEQ ID NO: 2:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 2818 amino acids  
TYPE: amino acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: cDNA to mRNA  
HYPOTHETICAL: NO  
ANTI-SENSE: NO  
ORIGINAL SOURCE:  
ORGANISM: Homo sapiens  
POSITION IN GENOME:  
CHROMOSOME/SEGMENT: 17q11.2  
FEATURE:  
NAME/KEY: Cleavage-site  
LOCATION: group(583..586, 815..818, 2573..2576, 2810..2813)  
OTHER INFORMATION: /note= "Potential CAMP-dependent  
OTHER INFORMATION: protein kinase recognition sites"  
FEATURE:  
NAME/KEY: Modified-site  
LOCATION: 2549..2556  
OTHER INFORMATION: /note= "Potential tyrosine  
OTHER INFORMATION: phosphorylation site"  
FEATURE:  
NAME/KEY: Modified-site  
LOCATION: group(1264, 1276, 1358, 1377, 1389, 1390, 1391,  
1395, 1396, 1400, 1423, 1426, 1429, 1430)  
OTHER INFORMATION: /note= "Invariant residues within  
OTHER INFORMATION: most statistically significant regions of similarity among th

OTHER INFORMATION: GAP family of proteins"  
FEATURE:  
NAME/KEY: Modified-site  
LOCATION: group(1264..1290, 1345..1407, 1415..1430)  
OTHER INFORMATION: /note= "Most statistically  
OTHER INFORMATION: significant regions of similarity among the GAP family of  
OTHER INFORMATION: proteins"  
FEATURE:  
NAME/KEY: Modified-site  
LOCATION: 496  
OTHER INFORMATION: /note= "At variance with previously  
OTHER INFORMATION: published sequence which shows an ATG methionine codon rat  
OTHER INFORMATION: than an ATA isoleucine codon"  
FEATURE:  
NAME/KEY: Modified-site  
LOCATION: 1183  
OTHER INFORMATION: /note= "At variance with previously  
OTHER INFORMATION: published sequence. Shows an CTG leucine codon rather than  
OTHER INFORMATION: previously published CTC"  
FEATURE:  
NAME/KEY: Modified-site  
LOCATION: 1555  
OTHER INFORMATION: /note= "At variance with previously  
OTHER INFORMATION: published sequence. Lacks an extra CAT histidine condon af  
OTHER INFORMATION: this residue"  
FEATURE:  
NAME/KEY: Modified-site  
LOCATION: (2771^2772)  
OTHER INFORMATION: /note= "Position of an 18 amino  
OTHER INFORMATION: acid insertion(SEQ ID NO:10) representing an alternatively  
OTHER INFORMATION: spliced product"  
FEATURE:  
NAME/KEY: Modified-site  
LOCATION: (1370^1371)  
OTHER INFORMATION: /note= "Position of a 21 amino acid  
OTHER INFORMATION: insertion representing an alternatively spliced product"  
FEATURE:  
NAME/KEY: Domain  
LOCATION: 1125..1537  
OTHER INFORMATION: /note= "NF1 catalytic domain"  
FEATURE:  
NAME/KEY: Modified-site  
LOCATION: 2746..2818  
OTHER INFORMATION: /note= "Corresponding amino acids  
OTHER INFORMATION: for the PstII-HindIII fragment designated pMAL.B3A"  
FEATURE:  
NAME/KEY: Modified-site  
LOCATION: 65..371  
OTHER INFORMATION: /note= "Corresponding amino acids  
OTHER INFORMATION: for the HpaI-PstI fragment designated pMAL.HF3A.p"  
FEATURE:  
NAME/KEY: Modified-site  
LOCATION: 65..1240  
OTHER INFORMATION: /note= "Corresponding amino acids  
OTHER INFORMATION: for the HpaI-XhoI fragment designated pMAL.HF3A.X"  
PUBLICATION INFORMATION:  
AUTHORS: Wallace, M.R. et al.  
TITLE: Type 1 Neurofibromatosis Gene: Correction  
JOURNAL: Science  
VOLUME: 250  
ISSUE: 12/21/90  
PAGES: 1749-  
DATE: 12/21-1990  
RELEVANT RESIDUES IN SEQ ID NO: 2: FROM 1 TO 2818  
PUBLICATION INFORMATION:  
AUTHORS: Wallace, M.R. et al.  
TITLE: Type 1 Neurofibromatosis Gene: Identification  
TITLE: of a large transcript in three NFL Patients  
JOURNAL: Science  
VOLUME: 249  
ISSUE: 07/13/90  
PAGES: 181-186  
DATE: 07/13-1990



RELEVANT RESIDUES IN SEQ ID NO: 2: FROM 1 TO 2818  
US-07-966-049A-2

Query Match 5.7% Score 100.5; DB 4; Length 2818;  
Best Local Similarity 22.2%; Pred. No. 1.3;  
Matches 86; Conservative 60; Mismatches 140; Indels 101; Gaps 18;  
QY 13 GVLGNEISLKSPGVSFRNGNWPIDG-----RIPDVAALSMGFSVKEDLSWPGGLAV 65  
DB 2000 GGLGSIKAEYMDATVALASGNKLVSSKVIKGMCKIIDTKLSPPTLEQHLMMDDIAI 2059  
QY 66 GNLHRRPRATVVMYKGNKALPPG-----SVISYPLENAVPEFSDS-----VANSTHS 115  
DB 2060 -----LARYMLSLFNNSLDVAHLPYLFHVTF-LVATGPLSLRSTHGLVNIHS 2111  
QY 116 L-----FSEETPVVQLAPSE-----ERYVMGKANSVFEDLSVTLR-OLNRNLFQENSV 164  
DB 2112 LCTCSQLHFSEETKQVRLSLTFESLPKFYLLFGISKVSAVIAFRSSYRDRSPGSGY 2171  
QY 165 -LSSLPLNSLRNNEVDLLFLSELQVLDHISL--LSRHKHLAK-----DHSPLDLYSLELA 217  
DB 2172 BRETFALTSLVTEA-LLEIME-ACMRDIPCTCKWLDQWTELAQRFAFOYNFSLQPRALV 2229  
QY 218 GLDRIKGR-YGEDESEQFRDASKIIVDALQKFAADMYS-----253  
DB 2230 VFGCISKRVSHGQIKOIRILSKALESCLK--GPDYNSQVLIETVIALTKLQPLLNKD 2287  
QY 254 -----LYGGNAVVEL--VTVKSFDTSLIRKTRTILEAKQAKNPASP-----292  
DB 2288 SPLKALFWAVAVQLDEVNLYSAGTALLEQNHLTDLRLFNDSKPEEVFMATRNPLE 2347  
QY 293 -----YNLAYKNFEYSVFNWV 310  
DB 2348 WHCKQMDHFVGLNFSNFNFALVGHLL 2374

RESULT 9  
US-08-751-189-3  
; Sequence 3, Application US/08751189  
; Patent No. 5919656  
; GENERAL INFORMATION:  
; APPLICANT: Harrington, Lea A.  
; APPLICANT: Robinson, Murray O.  
; TITLE OF INVENTION: No. 5919656el Genes Encoding Telomerase Protein  
; TITLE OF INVENTION: 1  
; NUMBER OF SEQUENCES: 12  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Amgen, Inc.  
; STREET: 1840 De Havilland Drive  
; CITY: Thousand Oaks  
; STATE: California  
; COUNTRY: USA  
; ZIP: 91320-1789  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: PatentIn Release #1.0, Version #1.30  
; CURRENT APPLICATION DATA:  
; FILING DATE: 15-NOV-1996  
; CLASSIFICATION: 435  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Oleski, Nancy A.  
; REGISTRATION NUMBER: 34,688  
; REFERENCE/DOCKET NUMBER: A-433  
; INFORMATION FOR SEQ ID NO: 3:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 2627 amino acids  
; TYPE: amino acid  
; STRANDEDNESS: unknown  
; TOPOLOGY: unknown

MOLECULE TYPE: protein  
US-08-751-189-3

Query Match 5.3% Score 93; DB 2; Length 2627;  
Best Local Similarity 22.5%; Pred. No. 7.1;  
Matches 66; Conservative 37; Mismatches 112; Indels 78; Gaps 10;  
QY 22 LKSPGSV--VFNGNWPIDGRIPIVDAALSMGFSVKEDLSWPGGLAVGNLFHRPRATVVM 79  
DB 1226 LKEPGALPSTYRSLSVWELQQLLPKSA-----ESL-----HPGQTQVLI 1264  
QY 80 VKGNKALPPGVSISYPLENAVPE-----FSLDSVANSTHSLFSEETPVVQLAPSEER 133  
DB 1265 IDGADRLVDOONQOLISDWIPKLPKRCVHLVLSVSSDAGLGETLEQSGAHVLAGPLEAS 1324  
QY 134 VYVMGKANSVFEDLSVTLRQLRNRLFOENSVL-----SSLPLNSLSRNNEVDLLFLSE- 186  
DB 1325 ----ARALVREELALYKGRLESPFNQMRLLLVKRESGRPLYLRVLTDLHLRLFTLYEQ 1380  
QY 187 ----LOVLDHISLSLRH--KHLAKDHSPLDLYSLELAGLOEIGKRYGDESEQFRDASKIL 240  
DB 1381 VSERLTLPATVPLLLQHLSTLEKEHGPDLVLPQALTAL-EVTR-----1423  
QY 241 VDLOKFAADMYSLYGGNAVVELVTVKSFDTSLIRKTRTILEAKQAKNPASP 293  
DB 1424 ----SGLTVDQLHGVLVSVWRTLPGTKSMEEAAAGNSGDPY 1461

RESULT 10  
US-09-060-836-3  
; Sequence 3, Application US/09060836  
; Patent No. 5981707  
; GENERAL INFORMATION:  
; APPLICANT: Harrington, Lea A.  
; APPLICANT: Robinson, Murray O.  
; TITLE OF INVENTION: No. 5981707el Genes Encoding Telomerase Protein  
; TITLE OF INVENTION: 1  
; NUMBER OF SEQUENCES: 12  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Amgen, Inc.  
; STREET: 1840 De Havilland Drive  
; CITY: Thousand Oaks  
; STATE: California  
; COUNTRY: USA  
; ZIP: 91320-1789  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: PatentIn Release #1.0, Version #1.30  
; CURRENT APPLICATION DATA:  
; FILING DATE: 15-NOV-1996  
; CLASSIFICATION:  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US 08/751,189  
; FILING DATE: 15-NOV-1996  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Oleski, Nancy A.  
; REGISTRATION NUMBER: 34,688  
; REFERENCE/DOCKET NUMBER: A-433  
; INFORMATION FOR SEQ ID NO: 3:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 2627 amino acids  
; TYPE: amino acid  
; STRANDEDNESS: unknown  
; TOPOLOGY: unknown  
; MOLECULE TYPE: protein  
US-09-060-836-3

Query Match 5.3% Score 93; DB 2; Length 2627;

```

Db 1226 LKEPGALPSTYRSLVWELQORLLPKSA-----ESL-----HPGQTQVLI 1264
QY 80 VKGVNKIALPPGQSVISPLENAVPP-----FSLDSVANSIHSLFSEETPVVLQAPSEER 133
Db 1265 IDGARLVDQNGQLISDWIPKLPVCRVFLVLSVSDAGLGETLEQSOCAHVLAIGPLEAS 1324
QY 134 VVMVGKANSVEFDLSVTLRNRLFOENSVL-----SSLPLNSLRNNEVDLLFLSE- 186
Db 1325 ----ARARLVREELALCYCKRLEESPNNQMRLLLVKRESCRPLYRLVTDHLELTLVEQ 1380
QY 187 ----LQVLHDJSLLSRH--KHLAKDHPDLYSLLELAGLDETKRYGSEDSEQFROASKIL 240
Db 1381 VSEKRLTPATVPLQLHLLSLLEKRBHGPDLVQALAL-EVTR----- 1423
QY 241 VDALQKPADDMYSLYGGNAVVELVTVKGSFDTSLIRKTRITILEAKOANKPASPY 293
Db 1424 -----SGLITVDQJHGVLSVWRTLPKCTKSWEAAVAGNSGDPY 1461

RESULT 12
US-09-143-954-4
Sequence 4. Application US/09143954A

```

```

? GENERAL INFORMATION:
? APPLICANT: Wallis, Nicola G.
? TITLE OF INVENTION: Murf
? FILE REFERENCE: GM10097
? CURRENT APPLICATION NUMBER: US/09/143,954A
? CURRENT FILING DATE: 1998-08-31
? EARLIER APPLICATION NUMBER: 60/060,011
? EARLIER FILING DATE: 1997-09-25
? NUMBER OF SEQ ID NOS: 6
? SOFTWARE: FastSeq for Windows Version 4.0
? SEQ ID NO 4:
? LENGTH: 448
? TYPE: ERT
? ORGANISM: Streptococcus pneumoniae
? US-09-143-954 -4

Query Match          5.2%; Score 91; DB 4; Length 448;
Best Local Similarity 22.5%; Pred. No. 0.61;
Matches 48; Conservative 39; Mismatches 66; Indels 60; gaps
QY      98 LENA VPFSLDSVANSISFLSEETPWVLQLAPSEERYVMVGKANSFEDLSVTLRNRR 157
        ||| : : | : : : : | : : : : | : : : : | : : : : | : : : : |

```

```

/ APPLICANT: Wallis, Nicola G.
/
/ TITLE OF INVENTION: MurF
/
/ FILE REFERENCE: GM10097
/
/ CURRENT APPLICATION NUMBER: US/09/143,954A
/
/ CURRENT FILING DATE: 1998-08-31
/
/ EARLIER APPLICATION NUMBER: 60/060,011
/
/ EARLIER FILING DATE: 1997-09-25
/

```



Tue Jul 3 16:27:28 2001

us-09-284-320-6.ra1

Page 10

Search completed: July 3, 2001, 16:18:53  
Job time: 165 sec

GenCore version 4.5  
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: July 3, 2001, 16:16:43 ; Search time 16.08 Seconds  
(without alignments)  
1658.028 Million cell updates/sec

Title: US-09-284-320-6

Perfect score: 1763

Sequence: 1 MAVFVLLALVAGVLGNEFS.....MDPGYDSIIYRMTNQKIRMD 350

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 219241 seqs, 76174552 residues

Total number of hits satisfying chosen parameters: 219241

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

PIR68:\*  
1: pir1.\*  
2: pir2.\*  
3: pir3.\*  
4: pir4.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Match	Length	ID	Description
1	1667	94.6	330	2 T08667	hypothetical prote
2	192.5	10.9	324	2 T23876	hypothetical prote
3	121.5	6.9	364	2 T52336	probable type 1 me
4	114	6.5	663	2 T40039	hypothetical prote
5	109.5	6.2	408	2 T25243	hypothetical prote
6	105	6.0	450	2 S35732	dnaA protein - Spi
7	105	6.0	458	2 T19941	hypothetical prote
8	104	5.9	412	2 G64685	hypothetical prote
9	103.5	5.9	502	2 F64543	conserved hypothet
10	103.5	5.9	732	2 T43619	secreted protein k
11	103.5	5.9	732	2 S30060	protein kinase ypk
12	103.5	5.9	925	2 C84538	probable LRR recep
13	103	5.8	470	2 A71685	guanosine pentapho
14	103	5.8	979	2 JQ0894	PII5 protein - Myc
15	102.5	5.8	1004	2 A39611	probable GTP-bindi
16	101	5.7	412	2 B71831	hypothetical prote
17	101	5.7	689	2 T40364	hypothetical prote
18	100.5	5.7	1358	2 A29360	hypothetical prote
19	100.5	5.7	2818	2 B55282	neurofibromatosis-
20	100.5	5.7	2820	2 JC5196	neurofibromin 1 -
21	99	5.6	524	2 J77851	probable aspartate
22	98.5	5.6	377	2 S21302	succinate dehydrog
23	98.5	5.6	471	2 T48743	probable 26S ATP/u
24	98.5	5.6	1038	1 MWRECB	myosin beta heavy
25	98.5	5.6	1935	1 A37102	myosin beta heavy
26	98	5.6	561	2 T75191	hypothetical prote
27	98	5.6	864	2 T08575	protein kinase hom
28	87.5	5.5	758	2 H75013	hypothetical prote
29	97.5	5.5	884	2 S73302	preprotein transio

30	97.5	5.5	2825	2 I54352	neurofibromin - mo
31	96.5	5.5	533	2 G72593	hypothetical prote
32	96.5	5.5	574	2 T51799	CIb1-like protein
33	96.5	5.5	1062	2 F83335	RND multidrug effl
34	96.5	5.5	1062	2 T30830	hypothetical prote
35	96.5	5.5	1934	2 I48153	myosin heavy chain
36	96.5	5.5	1938	1 S06003	myosin alpha heavy
37	96	5.4	390	2 E81122	bacteriophage DNA
38	95.5	5.4	1020	2 B82427	sensor protein for
39	95.5	5.4	1209	2 T21455	hypothetical prote
40	95.5	5.4	2471	2 T42977	large tegument pro
41	95	5.4	2253	2 T30336	nuclear/mitotic ap
42	94.5	5.4	1935	1 S06006	myosin beta heavy
43	94.5	5.4	1935	2 A59286	myosin heavy chain
44	94	5.3	1778	2 T50074	probable nucleopor
45	93.5	5.3	818	2 T31464	stage II sporulati

#### ALIGNMENTS

##### RESULT 1

T08667

hypothetical protein DKFp54700510.1 - human (fragment)

C:Species: Homo sapiens (man)

C:Date: 11-Jun-1999 #sequence\_revision 11-Jun-1999 #text\_change 23-Jul-1999

C:Accession: T08667

R:Poustka, A.; Klein, M.; Mewes, H.W.; Gassenhuber, J.; Wiemann, S.

submitted to the Protein Sequence Database, May 1999

A:Reference number: Z16466

A:Accession: T08667

A:Molecule type: mRNA

A:Residues: 1-330 <POU>

A:Cross-references: EMBL:AL049929

A:Experimental source: fetal brain; clone DKFp54700510

C:Genetics:

A:Note: DKFp54700510.1

Query Match 94.6%; Score 1667; DB 2; Length 330;  
Best Local Similarity 99.7%; Pred. No. 166; Indels 0; Gaps 0;  
Matches 329; Conservative 1; Mismatches 0;

QY	21	ILKSPGVVFRNGNPIPIGERIPDVAALSMGFSVKEDLSMPGLAVGNLFRPRATVMVMV	80
DB	1	ILKSPGVVFRNGNPIPIGERIPDVAALSMGFSVKEDLSMPGLAVGNLFRPRATVMVMV	60
QY	81	KGYNKALPPGVSIVSYPLENAVPFSDSVANSIHSLSFSEETPVVLOLAPSEERYVMGKA	140
DB	61	KGYNKALPPGVSIVSYPLENAVPFSDSVANSIHSLSFSEETPVVLOLAPSEERYVMGKA	120
QY	141	NSVFEDLSVTLRQLNRNLFQENSVLSLPLNSLSRNNEVDLLFSELQVLHDSLSLRH	200
DB	121	NSVFEDLSVTLRQLNRNLFQENSVLSLPLNSLSRNNEVDLLFSELQVLHDSLSLRH	180
QY	201	KHLAKHSDPLYSLEAGLDEIGKRYGDESEQFRDASKILVDALQKFAADMYSLYGGNAV	260
DB	181	KHLAKHSDPLYSLEAGLDEIGKRYGDESEQFRDASKILVDALQKFAADMYSLYGGNAV	240
QY	261	VELVTVKSDFTSLIRKTRTILEAKQAKNPASPNLAYKYNFYSVVFNNVLMIALALA	320
DB	241	VELVTVKSDFTSLIRKTRTILEAKQAKNPASPNLAYKYNFYSVVFNNVLMIALALA	300
QY	321	VIITSYNNWMDPGYDSIIYRMTNQKIRMD	350
DB	301	VIITSYNNWMDPGYDSIIYRMTNQKIRMD	330

##### RESULT 2

T23876

hypothetical protein R03E1.2 - Caenorhabditis elegans

C:Species: Caenorhabditis elegans

C:Date: 15-Oct-1999 #sequence\_revision 15-Oct-1999 #text\_change 18-Feb-2000

Db	6	VFWVALLLNTYRG-----EAGSGSYFFIDGSSNQYLPRSSSEALPKSPVEISAAYSA	59
QY	50	MGSFVKEDLSWPG-----LAVGNLRFPRATVMVMYGVNKIALPGSGVLSYPLENAV	103
Db	60	LGFAPSATLTADGSSKLNKILKNPFRPRAAFVLETAGADMLLTSPSHSF-LGNAIR	118
QY	104	FSLDSVANGTHLSFEETPVVQLAPSEERVVMYGVKANSFVDSVTLRQLRNKL----	158
Db	119	SSIKS-----DSYKADT----ELPDNEVVVVSVNEPSSDVTDKDI--NDFASWLGGSYV	166
QY	159	--FQENSVLSLPLNSLSRNEVDL-----LFLSELQVLHDISSLSRUKHLAK--DH	207
Db	167	AGAPDSGGIISLIPAL-CANVFENIKFAFKFATNIIIGLVNIPQVSVVDLISGIDP	225

[illegible]

```

000 QY      280 VLDERVNOESENLLFGSSRSSARSMEVEGIPSAIIA-----EVLIVRLTTLAWLTGII 334
      : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
000 Ddb      317 LALAVIITSYNIWNNDPGYDSIIYRMNOKI 347
      : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
000 QY      335 LLATILGIVYFLNMNPLTKDTLLY--SNVKL 363
      : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
      RESULT 4
      T40039
      hypothetical protein SPAC986.13 - fission yeast (Schizosaccharomyces pombe)
      C:Species: Schizosaccharomyces pombe
      C:Date: 03-Dec-1999 #sequence_revision 03-Dec-1999 #text_change 31-Jan-2000
      C:Accession: T40039; T40794
      R:W:Wood, V.; Rajandream, M.A.; Barrell, B.G.; Skelton, J.; Churcher, C.M.
      Submitted to the EMBL Data Library, September 1998
      A:Reference number: Z21901
      A:Accession: T40039
      A:A:Status: preliminary; translated from GB/EMBL/DDBJ
      A:A:Molecule type: DNA
      A:A:Residues: 1-663 <W0>
      A:A:Cross-references: EMBL:AL031517; PIDN:CAA20649.1; GSPDB:GN00067; SPDB:SPB
      A:A:Experimental source: strain 972h; cosmid c28E12
      R:W:Wood, V.; Rajandream, M.A.; Barrell, B.G.; Volckaert, G.
      Submitted to the EMBL Data Library, May 1999
      A:Reference number: Z21875
      A:Accession: T40794

```

```

A:Molecule type: DNA
A:Residues: 1-663 <W02>
A:Cross-references: EMBL:AL049769; PIDN: CAB42374.1; GSPDB:GN00067; SPDB:SP
A:Experimental source: strain 972h-; cosmid c986
C:Genetics:
A:Gene: SPBC28E12.02; SPAC986.13
A:Map position: 2

Query Match          6.5%;   Score 114;   DB 2;   Length 663;
Best Local Similarity 21.5%;   Pred. No. 1,5;
Matches 62;   Conservative 43;   Mismatches 92;   Indels 92;   Gaps

13  GVUGNEF-SILKPGSVVFNNGNWPIGERIPDVAALSMGFSYKEDLSWPGLVAGNLFHR 71
   :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :
414  GVSSSLSSIVSSTGIVETNG-----IGEK-----MSFSPLKLLSIP----- 451
      :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :
72  P-----RATVMVMKGVNKILAIIPGASVTSYDLENAVDPSISDVANSTSTISFSEFTDVLQ 126
      :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :

```

452 PTEFVAQLAIIICMASGVEML-LKTRNGIEYFQGVNTVP IAMDRAKSIIFYKFGSQWQHILL 510  
 127 LAPSEERYVYVWKANSVFEDLSVTLR-QLRNRLFQENSVLSSLPNLSLRNNEVDLLFLS 185  
 511 EAPTQDQDFISGKNGKLDKVKQOQCRFWLKG-----DILFCP 548

QY 186 ELQVLHDISSLSRKHAKDHPDLSYSLSELAGL-----DEIGKR 225  
 Db 549 QSTSIPTV-----DIYSELERVIGKMTMLLEFPAPMHPYVPEITHKK 592  
 QY 226 -YGESEQPRDASKILDALOKFADDMYSLYGGNAVVELVTVKSFDTSL 273  
 Db 593 LIGRGEQIQRTK-LYNSYIEFTTPTTCYGHNVLR--TPSKFSENL 638

RESULT 5  
 T25243  
 hypothetical protein T24D5.2 - Caenorhabditis elegans  
 C:Species: Caenorhabditis elegans  
 C>Date: 15-Oct-1999 #sequence\_revision 15-Oct-1999 #text\_change 04-Mar-2000  
 C:Accession: T25243  
 R:Wilkinson, J.  
 submitted to the EMBL Data Library, November 1995  
 A:Reference number: Z20003  
 A:Accession: T25243  
 A:Status: preliminary; translated from GB/EMBL/DDBJ  
 A:Molecule type: DNA  
 A:Residues: 1-408 <WIL>  
 A:Cross-references: EMBL:Z68012; PIDN:CAA92020.1; GSPDB:GN00028; CESP:T24D5.2  
 A:Experimental source: clone T24D5  
 C:Genetics:  
 A:Gene: CESP:T24D5.2  
 A:Map position: X  
 A:Introns: 114/2; 134/3; 206/3; 309/3; 355/2  
 C:Superfamily: Caenorhabditis elegans hypothetical protein T24D5.2

Query Match 6.2%; Score 109.5; DB 2; Length 408;  
 Best Local Similarity 20.8%; Pred. No. 1.6;  
 Matches 43; Conservative 47; Mismatches 84; Indels 33; Gaps 7;  
 QY 81 KGVNK-LALPPGVSIVS-----YPLENAVFPFSDLSVANSIHSLS-FSEETPVVVLQAPSEE 132  
 Db 103 KANQFLKPGDGEVLAKYFQCHLAKNTTKYKTYRLDAYKMTFSEERDSDILIRSDGDS 162  
 QY 133 RYVWGVKANSVEDLSVTLRQLNRFLQENSVLSSLPNLSLRNEVDLLFSELOVLHD 192  
 Db 163 TVRMGISTNLFNRNMAIEELRNLSQNIYKSLFVTNRSQO-----MGENDLSKC 215  
 QY 193 ISSLSRKHAKDHPDLSYSLSELAGLDEIGK-----RYGE--DSEQPRDASKILVDA 243  
 Db 216 VSTIAENLTVMYDGFPPNIIEIVKLHRLMKNFTLMCSGEGYGNIMELEQVNAANLKLDG 275  
 QY 244 LQFADDMYSLYGGNAVVELVTVKSD 270  
 Db 276 LSKITDDQ-----LNSITATNFE 293

RESULT 6  
 S35732  
 dnaA protein - Spiroplasma citri  
 C:Species: Spiroplasma citri  
 C>Date: 13-Jan-1995 #sequence\_revision 13-Jan-1995 #text\_change 02-Feb-2001  
 C:Accession: S35732  
 R:Ye, F.; Laigret, F.; Bove, J.  
 submitted to the EMBL Data Library, December 1992  
 A:Description: Nucleotide sequence and genetic organization at the replication origin  
 A:Reference number: S35732  
 A:Accession: S35732  
 A:Status: preliminary  
 A:Molecule type: DNA  
 A:Residues: 1-450 <YEF>  
 A:Cross-references: EMBL:Z19108; MID:g49345; PIDN:CAA79521.1; PID:g49346  
 C:Genetics:  
 A:Genetic code: SGC3  
 C:Superfamily: replication initiation protein dnaA  
 C:Keywords: DNA binding; nucleotide binding; P-loop  
 F:146-153/Region: nucleotide-binding motif A (P-loop)  
 F:208-212/Region: nucleotide-binding motif B

Query Match 6.0%; Score 105; DB 2; Length 450;  
 Best Local Similarity 20.4%; Pred. No. 3.9;  
 Matches 67; Conservative 41; Mismatches 115; Indels 106; Gaps 13;  
 QY 81 KGVNKIALPPGVSIVS-----YPLENAVFPFSDLSVANSIHSLSFSEETPVV 124  
 Db 83 KNINKQA-----SVISKIDILITENNLAYENYTFENFVR-----GDSNHEAMQALAVA 132  
 QY 125 LQAPSEERVYVWGKANSVEDLSVTLRQLNRFLQENS-----VLSLPNLSLRN 176  
 Db 133 LDLGKKWPLFIYDGLSKTHLLHAIENKVENIYKTNRRKYLKADEFGRKIAMDILNOG 192  
 QY 177 NEV-----DLLFISELQVL-----HDISLSSLRKH--LAKDHSPLY 212  
 Db 193 HEIEAFKTSYDIYDCLLIDDLQAKRNTNELFFHFNYSYKRNKQIVITSDKYPD-- 250  
 QY 213 SLELAGL-----EIGKRYGDSQPRDASKILVDAQ-----KF 247  
 Db 251 --DLGGFEARIISRFSGYGLSDSPDETALKILEQKLKHQNNLGLFSESELEFTALNF 308  
 QY 248 ADMYSLYGG-----NAVVELTV--KSFDTSLIRKTRTILEAKQAKNPASP 292  
 Db 309 NSDVRLKGAIKRLLFLAVMKNKPNIEITLADVEKAFKNAPLQNNKEITPKKIQIVADS 368  
 QY 293 YNLAYKYNFEYSVFNVMVIMITALAV 321  
 Db 369 YNITIKAMSKSRVSN---VMOARQLAM 393

RESULT 7  
 T19941  
 hypothetical protein C44H4.1 - Caenorhabditis elegans  
 C:Species: Caenorhabditis elegans  
 C>Date: 15-Oct-1999 #sequence\_revision 15-Oct-1999 #text\_change 18-Feb-2000  
 C:Accession: T19941  
 R:Smyle, R.  
 submitted to the EMBL Data Library, August 1996  
 A:Reference number: Z19200  
 A:Accession: T19941  
 A:Status: preliminary; translated from GB/EMBL/DDBJ  
 A:Molecule type: DNA  
 A:Residues: 1-458 <WIL>  
 A:Cross-references: EMBL:Z79598; PIDN:CAB01867.1; GSPDB:GN00028; CESP:C44H4.1  
 A:Experimental source: clone C44H4  
 C:Genetics:  
 A:Gene: CESP:C44H4.1  
 A:Map position: X  
 A:Introns: 24/3; 157/1; 257/3; 372/3; 425/1  
 Query Match 6.0%; Score 105; DB 2; Length 458;  
 Best Local Similarity 22.5%; Pred. No. 4;  
 Matches 69; Conservative 47; Mismatches 116; Indels 74; Gaps 12;  
 QY 18 EFSILKSPGVVFRNGNPIPERIPDVAALSGFSGKEDLSWPLGAVGNLFHRPRATVM 77  
 Db 149 QLKTLSPDNLi-----EYVDNAFLS-----YHSRSLK 179  
 QY 78 VMWGVNKLALPGSVISYPLENAVFPFSDLSVANSIHSLSFSEETPVVQLAPSEERVY 136  
 Db 180 LDLSANNLTAIHTGLLG--LENLSQLSLDKNLLSEIPSOALENIPSEDLISLGVNRIHT 237  
 QY 137 VGKANSV---FEDLSVTLRQLR---NRLFOENSVLSSLPi--NSLSRNEVDLLFSEL 187  
 Db 238 ISR-NSLPLNKLKSLLEVNOIRLIPSDSFSETPLLSYLGNLLTSIDASFMLHIGGL 296  
 QY 188 QVL-----HDISLSSLRKH-----HLAKDHSPLYSLSELAGLDEIGKRYGDS 231  
 Db 297 KVLMSNNKDDITSIQANGKLSVQOQFKICTICILAFQHPALSRLLEDFC----- 344  
 QY 232 QFRDASKILVDAQFADDMYSLYGGNAVVELVTVKSFDTSLIRKTRTILEAKQAKNPAS 291

1

Downloaded from <http://ajphaphysocpharm.sagepub.com/> at 11:51 11 November 2014

Downloaded from <http://ajphaphysocpharm.sagepub.com/> at 11:52 11 November 2014



Db 264 AGVVHNDI-----KPGNVFDRAS-----GE--PVVIDLGLHSRSGEQPKGFT--ESFKAPE 311  
 QY 63 LAVGNLFHRPRATYVMV-----KGVNKLALPPGVSIVSYPLENAVVP 103  
 Db 312 LGVGNLGAASEKSDVFLVSTLLHCIEGFEKPEKPNQGLRFTITSEPAHVM-----ENGYP 368  
 QY 104 FSLDSVANSIHLSEETPVVQLA-----PSEERVYMWGKANSVFEDLSVTLRQLRNR 157  
 Db 369 IHRPGIA-GVETAYTRFTIDILGVSADSRPDSNEARLHEFLSDGTIDESAKQI--LKDT 425  
 QY 158 LFQENSVLSPLNSLRNNEVDLLFLSELQVLHDISLSLRHKLAKHSDPL-----211  
 Db 426 LTGEMSPLS-----DVRRTIP-KKLRELSOLLRLTHLSAATKQOLDMGVLS 472  
 QY 212 YSLELAGLDEIGKRYGDESEQFRDASKILVDALOKFADDMWISLYGNAVVELVTKSFDT 271  
 Db 473 LDTMLVALDKAEREGGVKQD-----LKSFN- 498  
 QY 272 SLIRKTRTILE-----AKQAKNPASPNLAYKYNFEYSVV 306  
 Db 499 SLILKTYRVIEDYVKGREGDTKNSSTEVSPY---HRSNFMLSIV 539  
 RESULT 11  
 S30060  
 protein kinase ypkA - Yersinia pseudotuberculosis  
 C:Species: Yersinia pseudotuberculosis  
 C>Date: 06-Jan-1995 #sequence\_revision 06-Jan-1995 #text\_change 08-Oct-1999  
 C:Accession: S30060  
 R:Galyov, B.E.; Hakansson, S.; Forsberg, A.; Wolf-Watz, H.  
 Nature 361, 730-732, 1993  
 A:Title: A secreted protein kinase of Yersinia pseudotuberculosis is an indispensable virulence factor  
 A:Reference number: S30060; MUID:93180911  
 A:Accession: S30060  
 A:Status: preliminary  
 A:Molecule type: DNA  
 A:Residues: 1-732 <GAL>  
 A:Cross-references: EMBL:X69439; NID:g49186; PIDN:CAA49215.1; PID:g49187  
 C:Keywords: serine/threonine specific protein kinase

Query Match 5.9%; Score 103.5; DB 2; Length 732;  
 Best Local Similarity 21.2%; Pred. No. 10;  
 Matches 73; Conservative 51; Mismatches 103; Indels 117; Gaps 17;  
 QY 12 AGVLGNFSLKSPGVSFVRNGNWPPIGERIPDVAALSM-----GFSVKEDLSWPG 62  
 Db 264 AGVVHNDI-----KPGNVFDRAS-----GE--PVVIDLGLHSRSGEQPKGFT--ESFKAPE 311  
 QY 63 LAVGNLFHRPRATYVMV-----KGVNKLALPPGVSIVSYPLENAVVP 103  
 Db 312 LGVGNLGAASEKSDVFLVSTLLHCIEGFEKPEKPNQGLRFTITSEPAHVM-----ENGYP 368  
 QY 104 FSLDSVANSIHLSEETPVVQLA-----PSEERVYMWGKANSVFEDLSVTLRQLRNR 157  
 Db 369 IHRPGIA-GVETAYTRFTIDILGVSADSRPDSNEARLHEFLSDGTIDESAKQI--LKDT 425  
 QY 158 LFQENSVLSPLNSLRNNEVDLLFLSELQVLHDISLSLRHKLAKHSDPL-----211  
 Db 426 LTGEMSPLS-----DVRRTIP-KKLRELSOLLRLTHLSAATKQOLDMGVLS 472  
 QY 212 YSLELAGLDEIGKRYGDESEQFRDASKILVDALOKFADDMWISLYGNAVVELVTKSFDT 271  
 Db 473 LDTMLVALDKAEREGGVKQD-----LKSFN- 498

RESULT 12  
 C84538  
 probable LRR receptor protein kinase [imported] - Arabidopsis thaliana

C:Species: Arabidopsis thaliana (mouse-ear cress)  
 C>Date: 02-Feb-2001 #sequence\_revision 02-Feb-2001 #text\_change 02-Feb-2001  
 C:Accession: C84538  
 R:Lin, X.; Kaul, S.; Rounsley, S.D.; Shea, T.P.; Benito, M.I.; Town, C.D.; Fujii, C.Y.  
 M.; Koo, H.; Moffat, K.S.; Cronin, L.A.; Shen, M.; VanAken, S.E.; Umayam, L.; Tallon,  
 euss, D.; Nierman, W.C.; White, O.; Eisen, J.A.; Salzberg, S.L.; Fraser, C.M.; Venter  
 Nature 402, 761-768, 1999  
 A:Title: Sequence and analysis of chromosome 2 of the plant Arabidopsis thaliana.  
 A:Reference number: A84420; MUID:20083487  
 A:Accession: C84538  
 A:Status: preliminary  
 A:Molecule type: DNA  
 A:Residues: 1-925 <STO>  
 A:Cross-references: GB:AE002093; NID:g4544402; PIDN:AAD22312.1; GSPDB:GN00139  
 C:Genetics:  
 A:Gene: At2g16250  
 A:Map position: 2

Query Match 5.9%; Score 103.5; DB 2; Length 925;  
 Best Local Similarity 22.6%; Pred. No. 14;  
 Matches 86; Conservative 43; Mismatches 126; Indels 125; Gaps 16;

QY 4 FVYLL-----ALVAGVLGNEFS-----ILKSPGVVFRNGNWPPIGERIPD-----44  
 Db 11 FVILLCLVLPFDVGVVQOTSRFSEKILLINLSSGLRGTDWPIKGDPCVDNRGIOCE 70  
 QY 45 -----VAALSMGFSV-----KEDLSWPLAVGNLHFRPRAT 75  
 Db 71 NGSIIIGINISGFRTRIGKLNQFSDVPLRNLTLSYFNASGLALPG-TIPEWF-----123  
 QY 76 VMVWVGYNKALPPGVSIVSYPLENAVPSLDSVA-----NSIHLSEETPVVQL 126  
 Db 124 -----GVSLALEVLDDSSCVNGVVPFTIGNLTSTLTNLSONSLISPLSSGLQLLN 177  
 QY 127 LAPSE-ERYVMVGKANSVFEDLS--VTLRQLNRNL-----FOENSVL 165  
 Db 178 LSQDLRSNFTGVLPGQSFSLKNLLTLDVSSNYLTGPIPPGLGALSCLKLIHLNFSSNFS 237  
 QY 166 SSLP-----LNSLSRNEVDLLFLSELQVLHDISLSLRHKLAKHSDPD 210  
 Db 238 SPISELGLDVLNVDLFDLSINSLSGVPQELRKLKQLMAIGDNLSSGLTPV-----D 291  
 QY 211 LYSLELAGLDEIGKRYGDESEQFRDAS-----KILVDALOKFADDM--YSLYGGNAVVEL 263  
 Db 292 LFSAE-SQLQTLVLRENGFGSLPDCVWSLPKLRILDIKNNFTGLLPYSSYSDSDQIAEM 350  
 QY 264 VTKSPD-----TSLIRKTR 278  
 Db 351 VDSSNTFYGELTPIIRFR 370

RESULT 13  
 A71685  
 guanosine pentaphosphate phosphohydrolase (gppA) RP294 - Rickettsia prowazekii  
 C:Species: Rickettsia prowazekii  
 C>Date: 21-Nov-1998 #sequence\_revision 21-Nov-1998 #text\_change 03-Nov-2000  
 C:Accession: A71685  
 R:Andersson, S.G.E.; Zomorodipour, A.; Andersson, J.O.; Sicheritz-Ponten, T.; Alsmark  
 Nature 396, 133-140, 1998  
 A:Title: The genome sequence of Rickettsia prowazekii and the origin of mitochondria.  
 A:Reference number: A71630; MUID:99039499  
 A:Accession: A71685  
 A:Status: preliminary; nucleic acid sequence not shown; translation not shown  
 A:Molecule type: DNA  
 A:Residues: 1-470 <AND>  
 A:Cross-references: GB:AJ235271; GB:AJ235269; NID:g3868717; PIDN:CAA14755.1; PID:g386  
 A:Experimental source: strain Madrid E  
 C:Genetics:  
 A:Gene: gppA; RP294

Query Match 5.8%; Score 103; DB 2; Length 470;

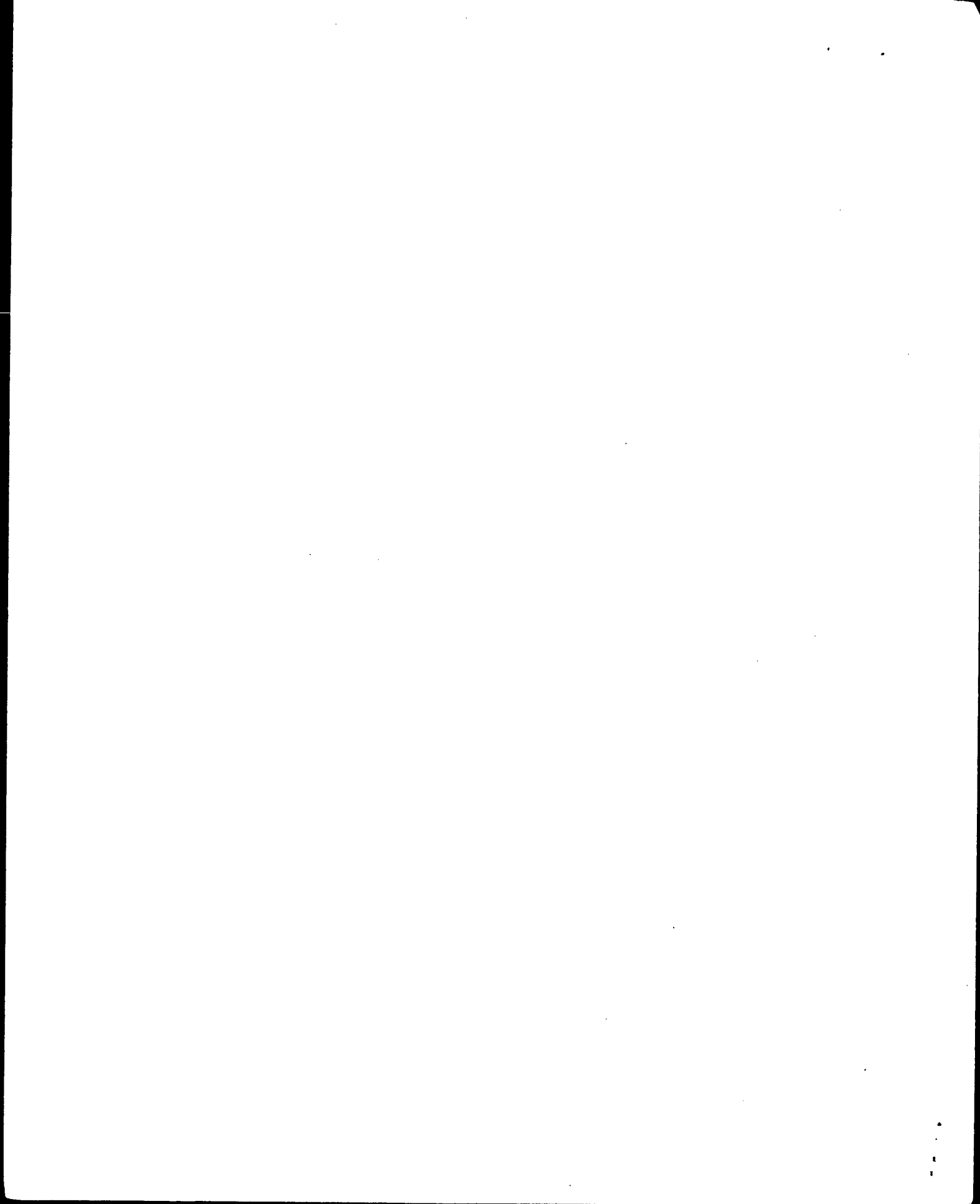


Tue Jul 3 16:27:28 2001

us-09-284-320-6.rpr

Page 7

---



GenCore version 4.5  
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: July 3, 2001, 16:17:23 ; Search time 12.1 Seconds  
(without alignments)  
990.861 Million cell updates/sec

Title: US-09-284-320-6  
Perfect score: 1763  
Sequence: 1 MAVFVLLALVAGVLGNEFS.....MDPGYDSIIYRMTNOKIRMD 350

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 93435 seqs, 34255486 residues

Total number of hits satisfying chosen parameters: 93435

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : SwissProt\_39:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match %	Length	DB ID	Description
1	513	29.1	100	1 VATN_HUMAN	O75787 homo sapien
2	122	6.9	30	1 VATN_BOVIN	P81134 bos taurus
3	105	6.0	450	1 DNAA_SPICI	P34028 spiroplasma
4	103.5	5.9	502	1 Y190_HELPY	P56117 helicobacte
5	103.5	5.9	732	1 YPKA_YERPS	Q05608 yersinia ps
6	103	5.8	979	1 P115_MYCHR	P41508 mycoplasma
7	102.5	5.8	1004	1 MV10_MOUSE	P23249 mus musculu
8	100.5	5.7	1358	1 SIR4_YEAST	P11978 saccharomyc
9	100.5	5.7	2839	1 NFL_HUMAN	P21359 homo sapien
10	99	5.6	483	1 SYD_MYCCA	Q48979 mycoplasma
11	98.5	5.6	1935	1 MYSB_HUMAN	P12883 homo sapien
12	97.5	5.5	884	1 SECA_FORPU	P51381 porphyra pu
13	97.5	5.5	2841	1 NFL_MOUSE	Q04690 mus musculu
14	97	5.5	715	1 ORC3_MOUSE	Q91330 mus musculu
15	96.5	5.5	805	1 SUS1_TULGE	Q41608 tulipa gesn
16	96.5	5.5	1934	1 MYSB_MESAU	P13540 mesocricetu
17	96.5	5.5	1938	1 MYSA_RAT	P02563 rattus norv
18	95.5	5.4	736	1 MYSB_RABIT	P04461 oryctolagus
19	94.5	5.4	1935	1 MYSB_PIG	P79293 sus scrofa
20	94.5	5.4	1935	1 MYSB_RAT	P02564 rattus norv
21	93.5	5.3	1939	1 MYSA_HUMAN	P13533 homo sapien
22	93	5.3	640	1 DXS_SYNY3	P73067 synecocyst
23	92.5	5.2	502	1 Y190_HELPY	Q92mp2 helicobacte
24	91.5	5.2	896	1 BOSS_DROME	P22815 drosophila
25	91.5	5.2	4092	1 DYHC_YEAST	P36022 saccharomyc
26	91	5.2	262	1 CUT8_SCHPO	P38937 schizosacch
27	91	5.2	427	1 FOLD_YEAST	Q12676 saccharomyc
28	90.5	5.1	465	1 MYSA_RABIT	P04460 oryctolagus
29	90.5	5.1	472	1 ATPB_RHOCA	P72247 rhodobacter
30	90.5	5.1	1102	1 MYSC_CHICK	P29616 gallus gall
31	90.5	5.1	1938	1 MYSA_MOUSE	Q02566 mus musculu
32	90	5.1	507	1 RECN_CAMJE	Q9phm5 campylobact
33	90	5.1	626	1 CE1B_ECOLI	P04479 escherichia

34 90 5.1 1032 1 KINN\_HUMAN O12840 homo sapien  
35 90 5.1 1085 1 CUT7\_SCHPO P24339 schizosacch  
36 89.5 5.1 642 1 DEAD\_KLEPN P33906 klebsiella  
37 89.5 5.1 968 1 Y682\_METJA Q58095 methanococc  
38 89.5 5.1 3106 1 LMA2\_MOUSE Q60675 mus musculu  
39 89 5.0 500 1 TACY\_CLOPE P19995 clostridium  
40 89 5.0 904 1 PMS1\_YEAST P14242 saccharomyc  
41 88.5 5.0 1663 1 CO3\_HUMAN P01024 homo sapien  
42 88 5.0 442 1 THDF\_MYCGE P47254 mycoplasma  
43 88 5.0 618 1 VEL\_HPV63 Q07847 human papil  
44 88 5.0 1177 1 MFD\_BACSU P37474 bacillus su  
45 87.5 5.0 505 1 HAS1\_YEAST Q03532 saccharomyc

## ALIGNMENTS

RESULT 1  
VATN\_HUMAN  
ID VATN\_HUMAN STANDARD; PRT; 100 AA.  
AC O75787;  
DT 30-MAY-2000 (Rel. 39, Created)  
DT 30-MAY-2000 (Rel. 39, Last sequence update)  
DE 01-OCT-2000 (Rel. 40, Last annotation update)  
DE VACUOLAR ATP SYNTHASE MEMBRANE SECTOR ASSOCIATED PROTEIN M8-9  
DE (V-ATPASE M8.9 SUBUNIT).  
OS Homo sapiens (Human).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.  
OX NCBI\_TaxID=9606;  
RN [1]  
RP SEQUENCE FROM N.A.  
RA Ludwig J., Schagger H., Boyd A., Appes D.K.;  
RL Submitted (AUG-1998) to the EMBL/GenBank/DBJ databases.  
CC -!- FUNCTION: VACUOLAR ATPASE IS RESPONSIBLE FOR ACIDIFYING A VARIETY  
CC OF INTRACELLULAR COMPARTMENTS IN EUKARYOTIC CELLS.  
CC -!- SUBUNIT: COMPOSED OF AT LEAST 10 SUBUNITS.  
CC -!- SUBCELLULAR LOCATION: INTEGRAL MEMBRANE PROTEIN (POTENTIAL).  
CC THIS SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
CC -----  
CC EMBL: Y17975; CAA76984.1; -  
CC KW Hydrolyase; Hydrogen ion transport; Transmembrane.  
CC FT TRANSMEM 53 73 POTENTIAL.  
CC SEQUENCE 100 AA; 11575 MW; A0705DB7E665F6DC CRC64;  
Query Match 29.1%; Score 513; DB 1; Length 100;  
Best Local Similarity 100.0%; Pred. No. 1.6e-31;  
Matches 100; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 251 MYSLYGNVAVELVTVKSFDTSLIRKTRILEAKQKNPASPYNLAYKYNFEYSVFNWV 310  
Db 1 MYSLYGNVAVELVTVKSFDTSLIRKTRILEAKQKNPASPYNLAYKYNFEYSVFNWV 60  
QY 311 LWIMIALALAVIITSYNNWMDPGYDSIIYRMTNOKIRMD 350  
Db 61 LWIMIALALAVIITSYNNWMDPGYDSIIYRMTNOKIRMD 100  
RESULT 2  
VATN\_BOVIN  
ID VATN\_BOVIN STANDARD; PRT; 30 AA.  
AC P81134;  
DT 15-DEC-1998 (Rel. 37, Created)  
DT 15-DEC-1998 (Rel. 37, Last sequence update)  
DE 30-MAY-2000 (Rel. 39, Last annotation update)



RA Cotton M.D., Weidman J.M., Fujii C., Bowman C., Watthey L., Wallin E.,  
 RA Hayes W.S., Borodovsky M., Karp D.D., Smith H.O., Fraser C.M.,  
 RA Venter J.C.;  
 RT "The complete genome sequence of the gastric pathogen *Helicobacter*  
 RT *pylori*.";  
 RL Nature 388:539-547(1997).  
 CC -|- SUBCELLULAR LOCATION: MEMBRANE-BOUND (POTENTIAL).  
 CC -|- SIMILARITY: BELONGS TO THE PHOSPHOLIPASE D FAMILY. CARDIOLIPIN  
 CC SYNTHASE SUBFAMILY. STRONG, TO E.COLI YMDC.  
 CC  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC  
 CC EMBL; AE000539; AAD07257.1;  
 DR TIGR; HP0190;  
 DR InterPro; IPR001736;  
 DR Pfam; PF00614; PLDc; 2;  
 KW Hypothetical protein; Transferase; Transmembrane.  
 FT TRANSMEM 1 21 POTENTIAL.  
 SQ SEQUENCE 502 AA; 58287 MW; C80911F2FB59C714 CRC64;

Query Match 5.9%; Score 103.5; DB 1; Length 502;  
 Best Local Similarity 22.9%; Pred. No. 2.5;  
 Matches 72; Conservative 35; Mismatches 115; Indels 93; Gaps 15;  
 QY 93 VISPLENA-----VPSLDSVANSHSLFSEETPVVQLAPSEEV-----YVMG 138  
 DB 95 VTAKELLNANRGVKVRILLDD--NGLDSDFD-----IMLNPHKNIEVKIFNPYYIRN 147  
 QY 139 KANSVFE---DLSVTQLRNLRFQENSVLSSPLNSLSRN---NEVDLFLFSLQVL-- 190  
 DB 148 KGLRYFEMADYERIKRHHNKLFIVDNFAVIGGRNIGDNYFDNDLDTNLF--DLDAFFF 206  
 QY 191 -----HDS--SLSRHHKLAHSDPLYSLEAGLDEIGKRYGESE 231  
 DB 207 GGVASKAKESFERYWRHRSIPVSLTRHKLNN-----AKEIAKLHEKIPISAEDKN 260  
 QY 232 QFRDASKILVDALQKFDADMYSLYGGNAVVELVTVKSFDTSLIRKTRTILEAKQAKNPAS 291  
 DB 261 QPEKKVNDFIDRFQKY---QYPIYVNAIFLADSPKKIDTPLYSPKIAFE--KALKNAKD 316  
 QY 292 PYNLAYKY-----NFEYSVVFNVLMIALALAVIITSYNINWMDPG 334  
 DB 317 SVFIASSYFIPGKMMKIFKNQISKIELNLTN-----SLSSTDAIVYGANE----- 365  
 QY 335 YDSIIRYMTNOKIRM 349  
 DB 366 -----RYRNQLVRM 374

RESULT 5  
 YPKA\_YERPS STANDARD; PRT; 732 AA.  
 AC Q05608;  
 DT 01-JUN-1994 (Rel. 29, Created)  
 DT 01-JUN-1994 (Rel. 29, Last sequence update)  
 DT 30-MAY-2000 (Rel. 39, Last annotation update)  
 DE PROTEIN KINASE YPKA PRECURSOR (EC 2.7.1.1).  
 GN YPKA  
 OS *Yersinia pseudotuberculosis*.  
 OG Plasmid pIB1.  
 OC Bacteria; Proteobacteria; gamma subdivision; Enterobacteriaceae;  
 OC *Yersinia*.  
 OX NCBI\_TaxID=633;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=YP111;

RX MEDLINE-93180911; PubMed-8441468;  
 RA Galyov E.E., Haakansson S., Forsberg A., Wolf-Watz H.;  
 RT "A secreted protein kinase of *Yersinia pseudotuberculosis* is an  
 RT indispensable virulence determinant.";  
 RL Nature 361:730-732(1993).  
 RN [2]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=YP111;  
 RX MEDLINE-94321324; PubMed-8045884;  
 RA Galyov E.E., Haakansson S., Wolf-Watz H.;  
 RT "Characterization of the operon encoding the YpkA Ser/Thr protein  
 RT kinase and the YopJ protein of *Yersinia pseudotuberculosis*.";  
 RL J. Bacteriol. 176:4543-4548(1994).  
 CC -|- FUNCTION: ACTS AS A VIRULENCE DETERMINANT.  
 CC -|- SUBCELLULAR LOCATION: SECRETED.  
 CC -|- SIMILARITY: BELONGS TO THE SER/THR FAMILY OF PROTEIN KINASES.  
 CC  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC  
 CC EMBL; X69439; CAA49215.1;  
 DR EMBL; L33833; AAA68487.1;  
 DR PIR; S30060; S30060.  
 DR InterPro; IPR000719;  
 DR InterPro; IPR002290;  
 DR PROSITE; PS00107; PROTEIN\_KINASE\_ATP; FALSE\_NEG.  
 DR PROSITE; PS00108; PROTEIN\_KINASE\_ST; 1  
 DR PROSITE; PS00111; PROTEIN\_KINASE\_DOM; 1  
 KW Virulence; Plasmid; transferase; Serine/threonine-protein kinase;  
 KW ATP-binding; Signal.  
 FT SIGNAL 1 ? POTENTIAL.  
 FT CHAIN 1 732 PROTEIN KINASE YPKA.  
 FT DOMAIN 136 408 PROTEIN KINASE.  
 FT ACT\_SITE 270 270 BY SIMILARITY.  
 SQ SEQUENCE 732 AA; 81707 MW; 7C09BDD4F9C9CBDF CRC64;

Query Match 5.9%; Score 103.5; DB 1; Length 732;  
 Best Local Similarity 21.2%; Pred. No. 4.2;  
 Matches 73; Conservative 51; Mismatches 103; Indels 117; Gaps 17;  
 QY 12 AGVLGNFSLKSPGVFRNGNWPPIGERIPDVAALSN-----GFSVKEDLSWPG 62  
 DB 264 AGVVHDI---KPGNVVFDRA---GE--PWVIDLGLHSRSGEQPKGFT--ESFKAPE 311  
 QY 63 LAVGNLHPRPRATVVMV-----KGVNKLALPGSVISVPLENAV 103  
 DB 312 LGVGNLGASEKSDVFLVWSTLLHCIEGFEKNPEIKPNOGLRFTITSEPAHYMD---ENGY 368  
 QY 104 FSLDSVANSHSLFSEETPVVQLA-----PSEERYVMVGKANSVFEDLSVTLRLNR 157  
 DB 369 IHRPGIA-GVETAYTRFTIDILGVSAADSRPDSNEARLHEFLSDGTIDEESAKQI--LKDT 425  
 QY 158 LPOENSVLSLPLNSLSRNNEVDLLFLSELOVLDHIFSLSRHKKHAKHSDPL----- 211  
 DB 426 LTGEMSPSLT-----DVRRTIP--KKURELSOLLTRHLSAAATKQDMDGVLS 472  
 QY 212 YSLEAGLDEIGKRYGESEQFRDASKILVDALQKFDADMYSLYGGNAVVELVTVKSFDT 271  
 DB 473 LDTMLVALDKAEREKGVKDKQ-----LKSFN- 498  
 QY 272 SLIRKTRTILE-----AKQAKNPASYPNLYKNFEYSV 306  
 DB 499 SLILKTYRVIEDYVKGREGDTKNSSTEVSPY----HRSNFMLSIV 539

RESULT 6  
 P115\_MYCHR

```

ID AC P115_MYCHR STANDARD; PRT; 979 AA.
DT P41508;
DT 01-NOV-1995 (Rel. 32, Created)
DT 01-FEB-1996 (Rel. 33, Last sequence update)
DT 15-JUL-1998 (Rel. 36, Last annotation update)
DE P115 PROTEIN.
OS Mycoplasma hyorhinis.
OC Bacteria; Firmicutes; Bacillus/Clostridium group; Mollicutes;
OC Mycoplasmataceae; Mycoplasma.
OX NCBI_TaxID=2100;
RN [1]
RP SEQUENCE FROM N.A. AND PARTIAL SEQUENCE.
RX MEDLINE=91138990; PubMed=1825306;
RA McArnica S.M., McIntosh M.A., Wise K.S.;
RT "A Mycoplasma hyorhinis protein with sequence similarities to
FT nucleotide-binding enzymes."
RL Gene 97:77-85(1991).
CC -|- SUBCELLULAR LOCATION: CYTOPLASMIC.
CC -|- DOMAIN: CONSISTS OF TWO PUTATIVE CENTRAL COILED-COIL REGIONS
CC FLANKED BY PUTATIVE GLOBULAR REGIONS AT THE N- AND C-TERMINUS.
CC -|- SIMILARITY: BELONGS TO THE SMC FAMILY. SIMILAR TO OTHER MYOPLASMA
CC P115.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announcement/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; M34956; AAA25423.1; -
DR PIR; JQ0894; JQ0894.
KW ATP-binding; Coiled coil.
FT NP_BIND 32 39
FT DOMAIN 169 224 COILED COIL (POTENTIAL).
FT DOMAIN 231 400 COILED COIL (POTENTIAL).
FT DOMAIN 569 821 COILED COIL (POTENTIAL).
FT DOMAIN 884 912 ALA/ASP-RICH (DA-BOX).
SQ SEQUENCE 979 AA; 110566 MW; 30D51C56B56280F4 CRC64;

Query Match 5.8%; Score 103; DB 1; Length 979;
Best Local Similarity 19.6%; Pred. No. 7;
Matches 67; Conservative 63; Mismatches 126; Indels 86; Gaps 15;

QY 11 VAGVLGNEFSLKSGSVVFRNGWNPGERIPDVAALSMGFSVKE-----D 57
DB 537 IAKLDHKYTVISLEGD-LFRPGGTITGSKLRTSILNYDIKIKETHTLTKFAEDQIHD 595
QY 58 LSWPGLAVGNLFRPRATVM-----VNVKGVNKLALPPGVSIVPLEN--A 101
DB 596 LKIKQOTIYNEIETVNSTIQVKIEANSKINILNEELANKLNASEIFKEQEDQES 655
QY 102 VPFSLDSVANSIHSLEETPVQLAPSEERYVW----CKANSVFEDLSVTLRQLNR 157
DB 656 LNLSPFSEKINIEKQISTLT---IELNSKDKRLTNLISEQKGKTKKQELDAKLRLNTQ 712
QY 158 LFOENSLSLPLNSLSRNEVDLLFLSELQVLHDISSLSRHKHLAKDHPDLSLELA 217
DB 713 -----HSDSITEQNAKFL-----VEQNRKLSLSEHYKLTLEAASEQYSLD-- 752
QY 218 GLDEIGKRYGDS--EQFRDASKILVDALQKADDMYSLYGNA--VVELTVKVS----- 268
DB 753 -LDIEQARHFVDSLKELKELGNVNEAETEP-EEVNVQRYQKQVIEELTTAKSKIEEA 810
QY 269 ---EDTSLIRKTKTILEAKQAKNPASPYNLAYKYNFEYSVVF 307
DB 811 ISLDEKIIINKITEIV-----NLV---NNEFNWVF 837

RESULT 7
MV10_MOUSE

```

```

ID MV10_MOUSE STANDARD; PRT; 1004 AA.
DT P23249;
DT 01-NOV-1991 (Rel. 20, Created)
DT 01-NOV-1991 (Rel. 20, Last sequence update)
DT 01-NOV-1997 (Rel. 35, Last annotation update)
DE PROTEIN MOV-10.
GN MOV10 OR GB110.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=91117255; PubMed=1899287;
RA Mooslechner K., Mueller U., Karls U., Hamann L., Harbers K.;
RT "Structure and expression of a gene encoding a putative GTP-binding
FT protein identified by provirus integration in a transgenic mouse
RT strain."
RL Mol. Cell. Biol. 11:886-893(1991).
RN [2]
RP SEQUENCE OF 1-45 FROM N.A.
RC STRAIN=BALB/C;
RX MEDLINE=94344088; PubMed=8065313;
RA Hamann L., Bayer K.U., Jensen K., Harbers K.;
RT "Interaction of several related GC-box- and Gp-box-binding proteins
RT with the intronic enhancer is required for differential expression of
RT the gbl10 gene in embryonal carcinoma cells."
RL Mol. Cell. Biol. 14:5786-5793(1994).
CC -|- FUNCTION: IMPORTANT FUNCTION IN DEVELOPMENT AND/OR CONTROL
CC OF CELL PROLIFERATION.
CC -|- SIMILARITY: BELONGS TO THE DNA2/NAM7 HELICASE FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announcement/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; X52574; CAA36803.1; -
DR EMBL; X75819; CAA53453.1; -
DR PIR; A39611; A39611.
DR MGD; MGI:97054; Mov10.
DR InterPro; IPR000886; -
DR PROSITE; PS00014; ER_TARGET; UNKNOWN_1.
KW ATP-binding; Helicase; Developmental protein.
FT NP_BIND 525 532 ATP (BY SIMILARITY).
FT SITE 646 649 DRAG_BOX.
SQ SEQUENCE 1004 AA; 113601 MW; F0305D581846E758 CRC64;

Query Match 5.8%; Score 102.5; DB 1; Length 1004;
Best Local Similarity 27.9%; Pred. No. 7.9;
Matches 56; Conservative 19; Mismatches 67; Indels 59; Gaps 9;

QY 149 VTLRQLRNRLFOENSLSVSS-----LPLNSLSRNEVD--LFLSELQVLHDISS 195
DB 299 ILLRQLPTLLGSGSIFTAKEVAEIKAELETTLSRNYEVKRLLLHLELQWHDHDI-- 356
QY 196 LLRSRHKHLAK-----DHSPLYSLELAGLDEIGKRYGDSFOPRASKILVDALQKF 247
DB 357 ---RHYDLDSPVMTWPDVQDNPRLLTLEVPVGAESRPSVLRGDHLF---ALLSSETQ- 408
QY 248 ADOMYSLYGNAVVELTVTK-SPTDSLIRKTKTILEAKQAKNPASPYNLAYKYNFEYS-- 304
DB 409 -DPVTVKGFVHKVELDRVKLSSTLSRFDV-----GLTFKVNFTENRQ 453
QY 305 -----VVFNNVLWIMI 315
DB 454 PLRVQHRALELTGRWLWPM 474

```



```

RESULT 8
SIR4_YEAST
ID SIR4_YEAST STANDARD; PRT: 1358 AA.
AC P11978;
DT 01-OCT-1989 (Rel. 12, Created)
DT 01-OCT-1989 (Rel. 12, Last sequence update)
DE 15-JUL-1999 (Rel. 36, Last annotation update)
DE REGULATORY PROTEIN SIR4 (SILENT INFORMATION REGULATOR 4).
GN SIR4 OR STE9 OR ASD1 OR UTH2 OR YDR227W OR YD9934.12.
OS Saccharomyces cerevisiae (Baker's yeast).
OC Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
OC Saccharomycetales; Saccharomycetaceae; Saccharomycetes.
OX NCBI_TaxID=4932;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=86142836; PubMed=3325825;
RA Marshall M., Mahoney D., Rose A., Hicks J.B., Broach J.R.;
RT "Functional domains of SIR4, a gene required for position effect
  regulation in Saccharomyces cerevisiae."
RL Mol. Cell. Biol. 7:4441-4452(1987).
RN [2]
RP SEQUENCE FROM N.A.
RX MEDLINE=95192063; PubMed=7885847;
RA Davies C.J., Hutchison C.A. III;
RT "Insertion site specificity of the transposon Tn3."
RL Nucleic Acids Res. 23:507-514(1995).
RN [3]
RP SEQUENCE FROM N.A.
RX STRAIN=S288C / AB972;
RA Murphy L., Harris D., Bartell B.G., Rajandream M.A.;
RL Submitted (MAR-1995) to the EMBL/GenBank/DBJ databases.
CC -!- FUNCTION: THE PROTEINS SIR1 THROUGH SIR4 CONTROL THE EXPRESSION OF
  MATING TYPE GENES RESIDENT AT LOCI OF EITHER END OF CHROMOSOME
  III. SIR3 AND SIR4 ASSOCIATE WITH THE C-TERMINUS OF RAP1 TO FORM A
  DNA-BINDING COMPLEX THAT INITIATES THE REPRESSION AT THE HM LOCI
  AND TELOMERES.
CC -!- SUBUNIT: INTERACTS WITH RAP1 C-TERMINUS.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
  between the Swiss Institute of Bioinformatics and the EMBL outstation -
  the European Bioinformatics Institute. There are no restrictions on its
  use by non-profit institutions as long as its content is in no way
  modified and this statement is not removed. Usage by and for commercial
  entities requires a license agreement (See http://www.isb-sib.ch/announce/
  or send an email to license@isb-sib.ch).
CC -----
DR EMBL; M37249; AAA20881.1; -
DR EMBL; U13239; AAC33144.1; -
DR EMBL; A48612; CAAB8507.1; -
DR PIR; A29360; A29360.
DR SGD; S0002635; SIR4.
KW Transcription regulation; Repressor; DNA-binding; Nuclear protein;
FT DOMAIN 1277 1347 COILED COIL (POTENTIAL).
FT VARIANT 994 994 P -> L.
SQ SEQUENCE 1358 AA; 152061 MW; 9C698765964F094E CRC64;

Query Match 5.7%; Score 100.5; DB 1; Length 1358;
Best Local Similarity 24.1%; Pred. No. 17;
Matches 70; Conservative 43; Mismatches 92; Indels 85; Gaps 17;

QY 38 PCRIPDVAALS-----MGFSVKEDLSWPLGVLGNLFHPRATVMVMGVNKL 87
| : : : : : : : : : : : : : : : : : : : : : : : : : : : :
| : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB 23 PNDKIPEREKSNVTKPKPLFTFAKSNYSRPSTAIHTSPHPSD--VKPTSHKQLQ 79
| : : : : : : : : : : : : : : : : : : : : : : : : : : : :
| : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 88 LPPGSVTSYPIENAVPFS--LDSVANS--IHSLFSETPVVLQAPSEERVYMGKANSVFE 145
| : : : : : : : : : : : : : : : : : : : : : : : : : : : :
| : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB 80 QPKSPSLKKNYNSPFSNLEKISNKLKLSLRST-----SAGRI-----ESNPN 127
| : : : : : : : : : : : : : : : : : : : : : : : : : : : :
| : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 146 DLSTVTLRQLNRFLF-----QENSVLSS-----LPLNSLRNNEVDLLFLSELQVLHDS 194
| : : : : : : : : : : : : : : : : : : : : : : : : : : : :
| : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB 128 DASRSLASFQTAFSRHAQQQTSTFNKPVRTIVPISTQTNNS-----FLS-----GVK 177
| : : : : : : : : : : : : : : : : : : : : : : : : : : : :
| : : : : : : : : : : : : : : : : : : : : : : : : : : : :

```



Db 2021 GGLSIIKAENWADAVASGNVKNVSKVIGRMCKIIDKTCISPTTLEQHLMDWDIAI 2080  
QY 66 GNLFRRPRATVVMVGVKNKALPPG-----SVISYPLENAVPSFLDS-----VANSIHS 115  
Db 2081 -----LARYMLSFNNLDVAHPLPYLFHWTF-LVATGPLSLRASTHGLVINIHS 2132  
QY 116 L-----FSEETPVQVLAPSE-----ERYVYGVKANSYFEDLSVTLR-QLNRNLFQENSV 164  
Db 2133 LCTCSQLHSEETKQVRLSLTSEFSLPKFYLLFGISKVKSAAVTAFRSSYRDRKSPSGSY 2192  
QY 165 -LSSPLNLSRNEVDLFLSELQVLDHDISSL--LSRHKHLAK-----DHSPDLXSLA 217  
Db 2193 BRETFAITSLETVEA-LLEIME-ACMRDIPCTCKWLDQWTELAQREAFQYNPSLOPRALV 2250  
QY 218 GLDEIGKR--YGGDSFOPRDASKTLVDALOKFADDMYS----- 253  
Db 2251 VFGISIRVSHGQIKQIRILSKALESCLK--GPTYNSQVLIETAVIALTKQLPLNKD 2308  
QY 254 -----LYGNAVVEL--VTVKSFDTSLIRKTRTILEAKOAKNPASP----- 292  
Db 2309 SPLHKALFWAVAVLQDDEVNLYSAGTALLEQNLHTLDSLRIFNDKSPBEVFAIRNPLE 2368  
QY 293 -----YNLAKYNEYSWENMV 310  
Db 2369 WHCKQMDHFVGLNFNSNFALVGHLL 2395  
RESULT 10  
ID SYD\_MYCCA STANDARD; PRT; 483 AA.  
AC Q48979;  
DT 01-NOV-1997 (Rel. 35, Created)  
DT 01-NOV-1997 (Rel. 35, Last sequence update)  
DT 01-NOV-1997 (Rel. 35, Last annotation update)  
DE ASPARTYL-TRNA SYNTHETASE (EC 6.1.1.12) (ASPARTATE--TRNA LIGASE)  
DE (ASPRS) (FRAGMENT).  
GN ASPS.  
OS Mycoplasma capricolum.  
OC Bacteria; Firmicutes; Bacillus/Clostridium group; Mollicutes;  
OC Entomoplasmataceae.  
OX NCBI\_TaxID:2095;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN-ATCC 27343 / KID;  
RX MEDLINE=96059641; PubMed=7476192;  
RA Bork P., Ouzounis C., Casari G., Schneider R., Sander C.,  
Dolan M., Gilbert W., Gillevet P.M.;  
RT "Exploring the Mycoplasma capricolum genome: a minimal cell reveals  
its physiology".  
RL Mol. Microbiol. 16:955-967(1995).  
CC -!- CATALYTIC ACTIVITY: ATP + L-ASPARTATE + TRNA(ASP) = AMP +  
PYROPHOSPHATE + L-ASPARTYL-TRNA(ASP).  
CC -!- SUBUNIT: HOMODIMER (BY SIMILARITY).  
CC -!- SUBCELLULAR LOCATION: CYTOPLASMIC.  
CC -!- SIMILARITY: BELONGS TO CLASS-II AMINOACYL-TRNA SYNTHETASE FAMILY.  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
between the Swiss Institute of Bioinformatics and the EMBL outstation -  
the European Bioinformatics Institute. There are no restrictions on its  
use by non-profit institutions as long as its content is in no way  
modified and this statement is not removed. Usage by and for commercial  
entities requires a license agreement (See <http://www.isb-sib.ch/announcement/>  
or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
CC -----  
DR EMBL; 233048; CA83720.1; -.  
DR InterPro; IPR002106; -.  
DR InterPro; IPR002309; -.  
DR Pfam; PF00152; TRNA-synt\_2; 1.  
DR PROSITE; PS00179; AA\_TRNA\_LIGASE II 1; 1.  
DR PROSITE; PS00339; AA\_TRNA\_LIGASE II 2; PARTIAL.  
KW Aminoacyl-trna synthetase; Protein biosynthesis; Ligase; ATP-binding.  
FT NON\_TER 1  
SQ SEQUENCE 483 AA; 56298 MW; A03F5BOCEA31087E CRC64;

Query Match 5.6%; Score 99; DB 1; Length 483;  
Best Local Similarity 25.8%; Pred. No. 5;  
Matches 62; Conservative 38; Mismatches 84; Indels 56; Gaps 14;  
QY 76 VMMVGVKNKALPPGCVISYPLENAVPSFLDSVANSIHSFSEETPVQVLAPSEERYV 135  
Db 31 IVKRVKSVNK-----ELITGEIEIVIKDOLL--VINK-----SELPFFVLE----- 68  
QY 136 MVGKANSYFEDLSVTLR--QLNRNLFQENSVLSSPLNLSRNEVDLFLSELQVLDH- 192  
Db 69 ---NDVNVNEDRLTYRLVLDLRQVWQCNLIIRA-KINHIRNVLTDLNL-EVETPYFA 123  
QY 193 -----LSSPLNLSRNEVDLFLSELQVLDHDISSL--LSRHKHLAK-----DHSPDLXSLA 217  
Db 124 KSTPEGARFLVPSRLNKNRYALPQSPQLFKQLLMISGID-----RYQIVRCFRD-EDL 178  
QY 240 LVDAQKQFAD-DWYSLYGVNAVVELVTVKSFDTSLIRKTRTILEAKOAKNPASPYNLAYK 298  
Db 179 RIDRQPEFTQLDLEMSFATSE-----DVMQISSESLIKKILKEVKNFEIKPL--LRLSYK 231  
RESULT 11  
ID MYSB\_HUMAN STANDARD; PRT; 1935 AA.  
AC P12883; Q14904; Q16579;  
DT 01-OCT-1989 (Rel. 12, Created)  
DT 01-JUN-1994 (Rel. 29, Last sequence update)  
DT 01-JUN-2000 (Rel. 40, Last annotation update)  
DE MYOSIN HEAVY CHAIN, CARDIAC MUSCLE BETA ISOFORM.  
GN MYH7 OR MYHCB.  
OS Homo sapiens (Human).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
OX NCBI\_TaxID:9606;  
RN [1]  
RP SEQUENCE FROM N.A.  
RX MEDLINE=91065634; PubMed=2249844;  
RA Jaenicke T., Diederich K.W., Haas W., Schleich J., Lichter P.,  
Pfordt M., Bach A., Vosberg H.P.;  
RT "The complete sequence of the human beta-myosin heavy chain gene and  
a comparative analysis of its product".  
RL Genomics 8:194-206(1990).  
RN [2]  
RP SEQUENCE FROM N.A.  
RX MEDLINE=90301496; PubMed=2362820;  
RA Liew C.-C., Sole M.J., Yamauchi-Takahara K., Kellam B.,  
Anderson D.H., Lin L., Liew J.;  
RT "Complete sequence and organization of the human cardiac beta-myosin  
heavy chain gene".  
RL Nucleic Acids Res. 18:3647-3651(1990).  
RN [3]  
RP SEQUENCE OF 1-115 FROM N.A.  
RX MEDLINE=89264452; PubMed=2726733;  
RA Yamauchi-Takahara K., Sole M.J., Liew J., Ing D., Liew C.-C.;  
RT "Characterization of human cardiac myosin heavy chain genes".  
RL Proc. Natl. Acad. Sci. U.S.A. 86:3504-3508(1989).  
RN [4]  
RP ERRATUM.  
RA Yamauchi-Takahara K., Sole M.J., Liew J., Ing D., Liew C.-C.;  
RL Proc. Natl. Acad. Sci. U.S.A. 86:7416-7417(1989).  
RN [5]  
RP SEQUENCE OF 1310-1935 FROM N.A.  
RX MEDLINE=86176778; PubMed=2421254;  
RA Saez L., Leinwand L.A.;  
RT "Characterization of diverse forms of myosin heavy chain expressed in  
adult human skeletal muscle".  
RL Nucleic Acids Res. 14:2951-2969(1986).  
RN [6]  
RP REVISTIONS.  
RA Leinwand L.A.;  
RL Submitted (MAR-1988) to the EMBL/GenBank/DBJ databases.

RN [7] SEQUENCE OF 1410-1935 FROM N.A.  
 RX MEDLINE-88299163; PubMed-2969919;  
 RA Kurabayashi M., Tsuchimoto H., Komuro I., Takaku F., Yazaki Y.;  
 RT "Molecular cloning and characterization of human cardiac alpha- and  
 RT beta-form myosin heavy chain complementary DNA clones. Regulation of  
 RT expression during development and pressure overload in human  
 RT atrium.";  
 RL J. Clin. Invest. 82:524-531(1988).  
 RN [8]  
 RP SEQUENCE OF 785-1935 FROM N.A.  
 RC TISSUE-Skeletal muscle;  
 RX MEDLINE-90235862; PubMed-1691980;  
 RA Bober E., Buchberger-Seidl A., Braun T., Singh S., Goede H.W.,  
 RA Arnold H.H.;  
 RT "Identification of three developmentally controlled isoforms of human  
 RT myosin heavy chains.";  
 RL Eur. J. Biochem. 189:55-65(1990).  
 RN [9]  
 RP SEQUENCE OF 1393-1935 FROM N.A.  
 RX MEDLINE-87192738; PubMed-3032769;  
 RA Jandreski M.A., Liew C.-C.;  
 RT "Construction of a human ventricular cDNA library and  
 RT characterization of a beta myosin heavy chain cDNA clone.";  
 RL Hum. Genet. 76:47-53(1987).  
 RN [10]  
 RP REVIEW ON VARIANTS  
 RX MEDLINE-96039076; PubMed-8533830;  
 RA Aral S., Matsuoka R., Hirayama K., Sukurai H., Tamura M., Ozawa T.,  
 RA Kimura M., Inamura S.-I., Furutani Y., Joh-O K., Kawana M., Takao A.,  
 RA Hosoda S., Momma K.;  
 RT "Missense mutation of the beta-cardiac myosin heavy-chain gene in  
 RT hypertrophic cardiomyopathy.";  
 RL Am. J. Med. Genet. 58:267-276(1995).  
 RN [11]  
 RP VARIANTS CMH1 GLN-256 AND ARG-741.  
 RX MEDLINE-93248216; PubMed-8483915;  
 RA Fanaapazir L., Dalakas M.C., Cyran F., Cohn G., Epstein N.D.;  
 RT "Missense mutations in the beta-myosin heavy-chain gene cause central  
 RT core disease in hypertrophic cardiomyopathy.";  
 RL Proc. Natl. Acad. Sci. U.S.A. 90:3993-3997(1993).  
 RN [12]  
 RP VARIANTS CMH1 GLN-403.  
 RX MEDLINE-90367131; PubMed-1975517;  
 RA Geisterfer-Lowrance A.A.T., Kass S., Tanigawa G., Vosberg H.-P.,  
 RA McKenna W., Seidman C.E., Seidman J.G.;  
 RT "A molecular basis for familial hypertrophic cardiomyopathy: a beta  
 RT cardiac myosin heavy chain gene missense mutation.";  
 RL Cell 62:999-1006(1990).  
 RN [13]  
 RP VARIANTS CMH1 GLN-249; GLN-403; CYS-453 AND MET-606.  
 RX MEDLINE-92204193; PubMed-1552912;  
 RA Watkins H., Rosenzweig A., Hwang D.S., Levi T., McKenna W.,  
 RA Seidmann C.E., Seidmann J.G.;  
 RT "Characteristics and prognostic implications of myosin missense  
 RT mutations in familial hypertrophic cardiomyopathy.";  
 RL New Engl. J. Med. 326:1108-1114(1992).  
 RN [14]  
 RP VARIANTS CMH1 GLN-403; CYS-453; ARG-584 AND MET-606.  
 RX MEDLINE-94070863; PubMed-8250038;  
 RA Watkins H., Thierfelder L., Anan R., Jarcho J., Matsumori A.,  
 RA McKenna W., Seidman J.G., Seidman C.E.;  
 RT "Independent origin of identical beta cardiac myosin heavy-chain  
 RT mutations in hypertrophic cardiomyopathy.";  
 RL Am. J. Hum. Genet. 53:1180-1185(1993).  
 RN [15]  
 RP VARIANTS CMH1 GLN-403 AND VAL-908.  
 RX MEDLINE-92346810; PubMed-1638703;  
 RA Epstein N.D., Cohn G.M., Cyran F., Fanaapazir L.;  
 RT "Differences in clinical expression of hypertrophic cardiomyopathy  
 RT associated with two distinct mutations in the beta-myosin heavy chain  
 RT gene. A 908Leu-->Val mutation and a 403Arg-->Gln mutation.";  
 RL Circulation 86:345-352(1992).

RN [16]  
 RP VARIANTS CMH1 LEU-403 AND TRP-403.  
 RX MEDLINE-94075629; PubMed-8254035;  
 RA Dausse E., Komaeda M., Feller L., Dubourg O., Dufour C., Carrier L.,  
 RA Wisniewsky C., Becovic J., Hengstenberg C., Al-Mahdawi S.;  
 RT "Familial hypertrophic cardiomyopathy. Microsatellite haplotyping and  
 RT identification of a hot spot for mutations in the beta-myosin heavy  
 RT chain gene.";  
 RL J. Clin. Invest. 92:2807-2813(1993).  
 RN [17]  
 RP VARIANTS CMH1 TRP-403.  
 RX MEDLINE-94093568; PubMed-8268932;  
 RA Moolman J.C., Brink P.A., Corfield V.A.;  
 RT "Identification of a new missense mutation at Arg403, a CpG mutation  
 RT hotspot, in exon 13 of the beta-myosin heavy chain gene in  
 RT hypertrophic cardiomyopathy.";  
 RL Hum. Mol. Genet. 2:1731-1732(1993).  
 RN [18]  
 RP VARIANTS CMH1 ASN-615.  
 RX MEDLINE-93038688; PubMed-1417858;  
 RA Nishi H., Kimura A., Harada H., Tushima H., Sasazuki T.;  
 RT "Novel missense mutation in cardiac beta myosin heavy chain gene  
 RT found in a Japanese patient with hypertrophic cardiomyopathy.";  
 RL Biochem. Biophys. Res. Commun. 188:379-387(1992).  
 RN [19]  
 RP VARIANTS CMH1 GLY-778.  
 RX MEDLINE-93343938; PubMed-8343162;  
 RA Harada H., Kimura A., Nishi H., Sasazuki T., Tushima H.;  
 RT "A missense mutation of cardiac beta-myosin heavy chain gene linked  
 RT to familial hypertrophic cardiomyopathy in affected Japanese  
 RT families.";  
 RL Biochem. Biophys. Res. Commun. 194:791-798(1993).  
 RN [20]  
 RP VARIANTS CMH1 VAL-908.  
 RX MEDLINE-93168485; PubMed-8435239;  
 RA Al-Mahdawi S., Chamberlain S., Cleland J., Nihoyannopoulos P.,  
 RA Gilliland D., French J., Choudhury L., Williamson R., Oakley C.;  
 RT "Identification of a mutation in the beta cardiac myosin heavy chain  
 RT gene in a family with hypertrophic cardiomyopathy.";  
 RL Br. Heart J. 69:136-141(1993).  
 RN [21]  
 RP VARIANTS CMH1 TRP-719.  
 RX MEDLINE-95179132; PubMed-7874131;  
 RA Greve G., Bachinski L., Friedman D.L., Czernuzewicz G., Anan R.,  
 RA Towbin J., Seidman C.E., Roberts R.;  
 RT "Isolation of a de novo mutant myocardial beta MHC protein in a  
 RT pedigree with hypertrophic cardiomyopathy.";  
 RL Hum. Mol. Genet. 3:2073-2075(1994).  
 RN [22]  
 RP VARIANTS CMH1 CYS-513; ARG-716, AND TRP-719.  
 RX MEDLINE-94110336; PubMed-8282798;  
 RA Anan R., Greve G., Thierfelder L., Watkins H., McKenna W., Solomon S.,  
 RA Vecchio C., Shono H., Nakao S., Tanaka H., Mares A. Jr., Towbin J.A.,  
 RA Spirito P., Roberts R., Seidman J.G., Seidman C.E.;  
 RT "Prognostic implications of novel beta cardiac myosin heavy chain gene  
 RT mutations that cause familial hypertrophic cardiomyopathy.";  
 RL J. Clin. Invest. 93:280-285(1994).  
 RN [23]  
 RP VARIANTS CMH1 THR-797.  
 RX MEDLINE-96047159; PubMed-7581410;  
 RA Moolman J.C., Brink P.A., Corfield V.A.;  
 RT "Identification of a novel Ala797Thr mutation in exon 21 of the beta-  
 RT myosin heavy chain gene in hypertrophic cardiomyopathy.";  
 RL Hum. Mutat. 6:197-198(1995).  
 RN [24]  
 RP VARIANTS CMH1 CYS-453.  
 RX MEDLINE-96209901; PubMed-8655135;  
 RA Ko Y.-L., Chen J.-J., Tang T.-K., Cheng J.-J., Lin S.-Y., Liou Y.-C.,  
 RA Kuan P., Wu C.-W., Lien W.-P., Liew C.-C.;  
 RT "Malignant familial hypertrophic cardiomyopathy in a family with a  
 RT 453Arg-->Cys mutation in the beta-myosin heavy chain gene:  
 RT coexistence of sudden death and end-stage heart failure.";  
 RL Hum. Genet. 97:585-590(1996).

[25]  
 RN VARIANTS CMH1 THR-349 AND TRP-719.  
 RP MEDLINE=98204402; PubMed=9544842;  
 RA Jeschke B., Uhl K., Weist B., Schroder D., Meitinger T.,  
 RA Dohlemann C., Vosberg H.-P.;  
 RT "A high risk phenotype of hypertrophic cardiomyopathy associated with

Query Match 5.6%; Score 98.5; DB 1; Length 1935;  
 Best Local Similarity 24.7%; Pred. No. 40;  
 Matches 46; Conservative 39; Mismatches 76; Indels 25; Gaps 8;

QY 115 SLFSETPVVLQ-----LAPSEERYVMGKANSVFEDLSVTLRLRLRNFQENSVLSS 167  
 DB 879 SLQEKNDLQVQAEONLADAEERCDLIK-NKI--QLEAKVKEMNERLEDEEMNAE 935  
 QY 168 LPLSLRNEVDLLFLSELQVLDHDISSL-----LSRKHKLAKDHPDLSLELAGLDE 221  
 DB 936 LPAKKRKLDE-----CSELK--RDIDDLTLAKVEKEKHATENKYNLI-TEEMAGLDE 987  
 QY 222 ICKRYGDSQFDRASKILVDALQKFAADMYSLYGNVAVVELTVKSFDTSLIRKTRTIL 281  
 DB 988 IIAKLTREKALQEAHQALDLDLQAEEDKVTNLTAKVKLE--QQVDDLEGSLEQEKVVM 1046  
 QY 282 EAKQAK 287  
 DB 1047 DLERAK 1052

RESULT 12  
 SECA PORPU  
 ID SECA PORPU STANDARD; PRT; 884 AA.  
 AC P51381;  
 DT 01-OCT-1996 (Rel. 34, Created)  
 DT 01-OCT-1996 (Rel. 34, Last sequence update)  
 DT 01-NOV-1997 (Rel. 35, Last annotation update)  
 DE PREPROTEIN TRANSLOCASE SECA SUBUNIT.  
 GN SECA.  
 OS Porphyra purpurea.  
 OC Chloroplast.  
 OC Eukaryota; Rhodophyta; Bangiophyceae; Bangiales; Bangiaceae; Porphyra.  
 OX NCBI\_TaxID=2787;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN-AVONPORT;  
 RA Reith M.E., Munnholland J.;  
 RT "Complete nucleotide sequence of the Porphyra purpurea chloroplast genome."  
 RL Plant Mol. Biol. Rep. 13:333-335(1995).  
 CC -!- FUNCTION: INVOLVED IN PROTEIN EXPORT AND IMPORT. SECA HAS A CENTRAL ROLE IN COUPLING THE HYDROLYSIS OF ATP TO THE TRANSFER OF PRE-SECRETORY PROTEINS ACROSS THE MEMBRANE.  
 CC -!- SUBCELLULAR LOCATION: CHLOROPLAST.  
 CC -!- SIMILARITY: BELONGS TO THE SECA FAMILY.  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----  
 CC EMBL; U38804; AAC08267.1; -;  
 DR Mendel; 10360; PORPU;seca;1.  
 DR InterPro; IPR000185; -;  
 DR Pfam; PF01043; Seca\_protein; 1.  
 DR PRINTS; PR00906; SECA.  
 DR PROSITE; PS01312; SECA; 1.  
 KW Protein transport; ATP-binding; Chloroplast; Translocation; Transport.  
 FT NP\_BIND 98 105 ATP (BY SIMILARITY).  
 SQ SEQUENCE 884 AA; 101325 MW; A415846D12B90B2B CRC64;

Query Match 5.5%; Score 97.5; DB 1; Length 884;  
 Best Local Similarity 22.8%; Pred. No. 15;  
 Matches 61; Conservative 47; Mismatches 92; Indels 67; Gaps 13;

QY 94 ISVPLENAPFSLDSVANSI-----HSLFSEETPVVLQALPSEERYVMGKAN 141  
 DB 451 IPHSLNAPENVENKESDIIAAGROSSVTIATNAGRGTDIIILGNPS-----YI---AK 503  
 QY 142 SVFEDLSVTLRLRNL-FQENSVLSPSLNLSRNNEVDLLF--LSELQVLDHIS----- 194  
 DB 504 SILDDLRIKISVQNNKLOQLSLKTYCINQILKSLLEDLLIYANLSVLELEKKTISTACE 563  
 QY 195 -SLSRKHKLAKDHPDLSL-----YSLELAGLDEIGKRYGDE---SQEPR-- 234  
 DB 564 QVALSRNLETOLRKAYQLIFOEYENIFSQEKKYVAAGLGHVIGTERHESRRIDNQLRGR 623  
 QY 235 -----DASKILVDALQKFAADMYSLYGNVAVVELTVKSFDT-----TSLIRKTRIL 281  
 DB 624 AGRGDPGSSRFLL-----SIDNLLRIFGNGKIADLMQALNVDDTTPMESTLLSKS---L 676  
 QY 282 EAKQAKNPASPYNLAYKYNFEYSVFN 308  
 DB 677 EAAQKVEAYFYD-TRKQVFEYDQVLN 702

RESULT 13  
 NFL\_MOUSE  
 ID NFL\_MOUSE STANDARD; PRT; 2841 AA.  
 AC Q04690; O61956; O61957;  
 DT 01-JUN-1994 (Rel. 29, Created)  
 DT 01-JUN-1994 (Rel. 29, Last sequence update)  
 DT 30-MAY-2000 (Rel. 39, Last annotation update)  
 DE NEUROFIBROMIN (NEUROFIBROMATOSIS-RELATED PROTEIN NF-1).  
 GN NFL.  
 OS Mus musculus (Mouse).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
 OX NCBI\_TaxID=10090;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN-BALB/C; TISSUE-Brain;  
 RX MEDLINE=93357730; PubMed=8353485;  
 RA Bernards A., Snijders A.J., Hannigan G.E., Murthy A.E., Gusella J.F.;  
 RT "Mouse neurofibromatosis type 1 cDNA sequence reveals high degree of conservation of both coding and non-coding mRNA segments."  
 RL Hum. Mol. Genet. 2:645-650(1993).  
 RN [2]  
 RP SEQUENCE OF 1178-1555 FROM N.A., AND ALTERNATIVE SPLICING.  
 RX MEDLINE=95047432; PubMed=7958951;  
 RA Mantani A., Makasugi S., Yokota Y., Abe K., Ushio Y., Yamamura K.;  
 RT "A novel isoform of the neurofibromatosis type-1 mRNA and a switch of isoforms during murine cell differentiation and proliferation."  
 RL Gene 148:245-251(1994).  
 RN [3]  
 RP SEQUENCE OF 1950-2568 FROM N.A.  
 RX MEDLINE=90384569; PubMed=2169593;  
 RA Buchberg A.M., Cleveland L.S., Jenkins N.A., Copeland N.G.;  
 RT "Sequence homology shared by neurofibromatosis type-1 gene and IRA-1 and IRA-2 negative regulators of the RAS cyclic AMP pathway."  
 RL Nature 347:291-294(1990).  
 CC -!- FUNCTION: STIMULATES THE GTPASE ACTIVITY OF RAS. NFL SHOWS GREATER AFFINITY FOR RAS GAP, BUT LOWER SPECIFIC ACTIVITY. THUS IT MAY BE A REGULATOR OF RAS ACTIVITY.  
 CC -!- ALTERNATIVE PRODUCTS: 4 ISOFORMS; I, II (SHOWN HERE), III AND IV; ARE PRODUCED BY ALTERNATIVE SPLICING.  
 CC -!- TISSUE SPECIFICITY: TYPE I IS EXPRESSED PREDOMINANTLY IN BRAIN, SPINAL CORD AND TESTIS. TYPE II IS EXPRESSED PREDOMINANTLY IN ADRENAL GLAND, KIDNEY, OVARY AND LUNG. TYPE III IS EXPRESSED PREDOMINANTLY IN ADRENAL GLAND AND TYPE IV IS EXPRESSED MAINLY IN THE TESTIS.  
 CC -!- SIMILARITY: TO OTHER RAS GTPASE-ACTIVATING PROTEINS.  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration

CC the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----  
 CC EMBL; L10369; AAA39806.1; -  
 CC EMBL; L10367; AAA39806.1; JOINED.  
 CC EMBL; L10368; AAA39806.1; JOINED.  
 CC EMBL; L10370; AAA68132.1; -  
 CC EMBL; X54924; CAA38690.1; -  
 CC EMBL; D30730; BAA06395.1; -  
 CC EMBL; D30731; BAA06396.1; -  
 CC MGD; MGI:97306; Nf1.  
 CC InterPro; IPR001936; -  
 CC Pfam; PF00616; RasGAP; 1.  
 CC PROSITE; PS00509; RAS\_GTPASE\_ACTIV\_1; 1.  
 CC PROSITE; PS00118; RAS\_GTPASE\_ACTIV\_2; 1.  
 CC GTPase activation; Alternative splicing.  
 CC DOMAIN 1237 1453  
 CC RAS-GAP.  
 CC MISSING (IN ISOFORM I AND ISOFORM IV).  
 CC VVSORFQNSIGA -> VPKSSCFCLNNRWLASLRT  
 CC ASVP (IN ISOFORM III AND ISOFORM IV).  
 CC MISSING (IN ISOFORM III AND ISOFORM IV).  
 CC VARSPLIC 1373 1393  
 CC MISSING (IN ISOFORM III AND ISOFORM IV).  
 CC VARSPLIC 1394 1406  
 CC MISSING (IN ISOFORM III AND ISOFORM IV).  
 CC VARSPLIC 1407 2841  
 CC MISSING (IN ISOFORM III AND ISOFORM IV).  
 CC SEQUENCE 2841 AA; 319591 MW; 06650BDCDC531002 CRC64;

Query Match 5.5%; Score 97.5; DB 1; Length 2841;  
 Best Local Similarity 22.0%; Pred. No. 83;  
 Matches 85; Conservative 61; Mismatches 140; Indels 101; Gaps 18;

QY 13 GVLGNEFSILKSPGVYFRNGNWPGE-----RIPDVAALSMGFSEKEDLSWPGGLAV 65  
 DB 2023 GGLGSIKAEVMDTAVALASGNVGLVSKVIGRMCKIIDKCLSPPTLEQLHLMDDIAI 2082  
 QY 66 GNLFHRPRATVMVWGVNKLALPG-----SVISYPLENAVPPSLDS-----VANSTHS 115  
 DB 2083 -----LARYMLMFSNLDVAHAHLPYLFHVTF-LVATGPLSLRSTHGLNINIHS 2134  
 QY 116 L-----FSEETPVVQLAPSE-----ERYVMGKANSFEDLSVTLR-OLRNRLFOENSV 164  
 DB 2135 LCTCSQLHFSEETQVLRSLTFESLPKFLLEGISKVSAAVAFRSSVDRDSFSPGSY 2194  
 QY 165 -LSSPLNSLRNNEVDLFLSEQLVLDISSL--LSRHKLAK-----DHSPLYSLELA 217  
 DB 2195 ERETALTSLETYTEA-LLEIME-ACHRDIPCTCKWLDQWTELQAFQYFQYNSLOPRALV 2252  
 QY 218 GLDEIGKR--YGEDESEQFRDASKILVDALQKFADDMS-----253  
 DB 2253 VFGCISKRVSHGQIKIIRILSKALESCLK--GPDYNSQVLIESTVIALTKIQLPLLNKD 2310  
 QY 254 -----LYGNAVVEL--VTVKSFDTSLRKTRTILEAKQKNPASP-----292  
 DB 2311 SPLHKALFWAVAVLQDDEVNLYSNGTALLEQNLTLDLSLRFNDKSPFEVFWAIRNPLE 2370  
 QY 293 -----YNLAYKNFYSPVFNNA-310  
 DB 2371 WHCKOMDHFVGLAFNSNFALVGHLL 2397

RESULT 14  
 ORC3\_MOUSE  
 ID ORC3\_MOUSE STANDARD; PRT; 715 AA.  
 AC Q9JK30;  
 DT 01-OCT-2000 (Rel. 40, Created)  
 DT 01-OCT-2000 (Rel. 40, Last sequence update)  
 DT 01-OCT-2000 (Rel. 40, Last annotation update)  
 DE ORIGIN RECOGNITION COMPLEX SUBUNIT 3 (ORIGIN RECOGNITION COMPLEX  
 DE SUBUNIT LATHEO).  
 GN ORC3L OR ORC3.  
 OS Mus musculus (Mouse).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
 OX NCBI\_taxID=10090;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RX PubMed=10702681;  
 RA Springer J., Nanda I., Hoehn K., Schmid M., Grummt F.;  
 RT "Identification and chromosomal localization of murine ORC3, a new  
 member of the mouse origin recognition complex.";  
 RL Cytogenet. Cell Genet. 87:245-251(1999).  
 CC -!- FUNCTION: COMPONENT OF THE ORIGIN RECOGNITION COMPLEX (ORC) THAT  
 CC BINDS ORIGINS OF REPLICATION. IT HAS A ROLE IN BOTH CHROMOSOMAL  
 CC REPLICATION AND MATING TYPE TRANSCRIPTIONAL SILENCING. BINDS TO  
 CC THE ARS CONSENSUS SEQUENCE (ACS) OF ORIGINS OF REPLICATION IN AN  
 CC ATP-DEPENDENT MANNER.  
 CC -!- SUBUNIT: ORC IS COMPOSED OF SIX SUBUNITS (BY SIMILARITY).  
 CC -!- SUBCELLULAR LOCATION: NUCLEAR.  
 CC -!- SIMILARITY: BELONGS TO THE ORC3 FAMILY.  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----  
 CC EMBL; AJ132830; CAB76399.1; -  
 CC MGD; MGI:1394944; Orc3.  
 CC DNA replication; Nuclear protein; DNA-binding.  
 CC SEQUENCE 715 AA; 82341 MW; 1E915A719A2E33EB CRC64;

Query Match 5.5%; Score 97; DB 1; Length 715;  
 Best Local Similarity 22.0%; Pred. No. 12;  
 Matches 80; Conservative 45; Mismatches 116; Indels 122; Gaps 17;

QY 67 NLFHRPRATVMVWGVNKLAL-----PPGSVSYPLENAV-----FSLDSVANSHSL 116  
 DB 318 NIFLYHDFSISQSFIGI-KLSLEHFFYSQPLSVLCCLSEAKRVNVFSQCCN-IRRL 375  
 QY 117 -----FSEETPVVQLAPSEERYVMGKANSFEDLSV-----TLROLNRLFOENSV 164  
 DB 376 PSFRRYVENOPLKQVALLTNETFLKEKTQSLLEDLHYHYNFLVLRCLHN-----F 428  
 QY 165 LSSPLNSLRNNEVDLFLSEQLV-----LHDISLSLRHKLAKHSDPLYSLEAGLD 220  
 DB 429 TSSLPKPLGR--QIRELYCTCLEKTIWSEYKSAQLQLRLAKDELVSILQRCIEYLD 486  
 QY 221 -----ETG-----KRYGESEQFRDASKILVDALQKFADDMSYLYGNAVVE 262  
 DB 487 SSTEKQLGNTQIKIDFLTQFONLDADSKEEEDACGSPKGLQK--TDLYHLQ--KSILE 542  
 QY 263 L-----VTVKSFDTSLIR-----KTRTILEAKQA--286  
 DB 543 MKELURTKKPTKPEMLRENVMNFIIDLVRDYLLPPESQPLHEVVYFSAANTLREHNAAP 602  
 QY 287 -----KNPASPYNLAYKYNFYFVSVMVFNWLMIALALAVI 322  
 DB 603 RIALHTALNPPYYLKNEELEGCPINTAPDICIAYKLHLECSLLINLVDM---AFAPATV 659  
 QY 323 ITS 325  
 DB 660 VTA 662

RESULT 15  
 SUSL\_TULGE  
 ID SUSL\_TULGE STANDARD; PRT; 805 AA.  
 AC Q41608;  
 DT 15-DEC-1998 (Rel. 37, Created)  
 DT 15-DEC-1998 (Rel. 37, Last sequence update)  
 DT 15-DEC-1998 (Rel. 37, Last annotation update)  
 DE SUCROSE SYNTHASE 1 (EC 2.4.1.13) (SUCROSE-UDP GLUCOSYLTRANSFERASE 1).

```

OS Tulipa gesneriana.
OC Magnoliata: Viridiplantae; Embryophyta; Tracheophyta; Spermatophyta;
CC Eukaryotophyta; Liliopsida; Liliales; Liliaceae; Tulipa.
OX NCBI_TaxID=13306;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=CV. APELDOORN;
RA Baik P.A., de Boer A.D.;
RL Submitted (MAR-1996) to the EMBL/GenBank/DBJ databases.
RC -1- FUNCTION: SUCROSE-CLEAVING ENZYME THAT PROVIDES UDP-GLUCOSE AND
CC FRUCTOSE FOR VARIOUS METABOLIC PATHWAYS.
CC CATALYTIC ACTIVITY: UDP-GLUCOSE + D-FRUCTOSE = UDP + SUCROSE.
CC -1- SIMILARITY: TO OTHER PLANTS SUCROSE SYNTHASE.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL: X96938; CAA65639.1; -.
CC DR InterPro: IPR000368; -.
CC DR InterPro: IPR001296; -.
CC DR Pfam: PF00534; Glycos_transf_1; 1.
CC DR Pfam: PF00862; Sucrose_synth; 1.
CC DR Transference; Glycosyltransferase; Multigene family.
CC SKW SEQUENCE 805 AA; 92254 MW; B749D8953C9338F0 CRC64;

```

7

!



GenCore version 4.5  
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: July 3, 2001, 16:17:48 ; Search time 28.21 Seconds  
(without alignments)  
1641.502 Million cell updates/sec

Title: US-09-284-320-6

Perfect score: 1763

Sequence: 1 MAVFVLLALVAGVLGNEFS.....MDPGYDSIIYRINQKIRMD 350

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 425026 seqs, 132305027 residues

Total number of hits satisfying chosen parameters: 425026

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database :

SPTREMBL\_16:\*  
1: sp.archaea:\*  
2: sp.bacteria:\*  
3: sp.fungi:\*  
4: sp.human:\*  
5: sp.invertebrate:\*  
6: sp.mammal:\*  
7: sp.mhc:\*  
8: sp.organelle:\*  
9: sp.phage:\*  
10: sp.plant:\*  
11: sp.rodent:\*  
12: sp.unclassified:\*  
13: sp.vertebrate:\*  
14: sp.virus:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

# SUMMARIES

Result No.	Score	Match	Length	ID	Description
1	1759	99.8	350	4 Q9H2P8	Q9h2p8 homo sapien
2	1667	94.6	330	4 Q9UG89	Q9ug89 homo sapien
3	327	18.5	320	5 Q9VHG4	Q9vnh4 drosophila
4	192.5	10.9	324	5 Q9XTY9	Q9xt9 caenorhabdi
5	121.5	6.9	364	10 Q9ZSJ7	Q9zs37 arabidopsis
6	114	6.5	663	3 Q74359	Q74359 schizosacch
7	109.5	6.2	408	5 Q22736	Q22736 caenorhabdi
8	107.5	6.1	251	5 Q9N598	Q9n598 caenorhabdi
9	107	6.1	1112	10 Q9FH47	Q9fh47 arabidopsis
10	105	6.0	458	5 Q93377	Q93377 caenorhabdi
11	104	5.9	412	2 Q25885	Q25885 helicobacte
12	103.5	5.9	732	2 Q68717	Q68717 yersinia pe
13	103.5	5.9	732	2 Q9R112	Q9rl12 yersinia pe
14	103.5	5.9	925	10 Q9SIX4	Q9six4 arabidopsis
15	103	5.8	470	2 Q9ZDN3	Q9zdn3 rickettsia
16	102.5	5.8	1380	5 Q9VIU8	Q9vlu8 drosophila
17	102	5.8	403	4 Q9NX71	Q9nx71 homo sapien
18	101	5.7	412	2 Q9ZJQ7	Q9zjq7 helicobacte
19	101	5.7	689	3 P87169	P87169 schizosacch

20	100.5	5.7	2820	11	P97526	P97526 rattus norv
21	99.5	5.6	861	5	Q9N519	Q9n519 caenorhabdi
22	98.5	5.6	471	3	Q9P723	Q9p723 neurospora
23	98.5	5.6	728	2	Q56921	Q56921 yersinia en
24	98.5	5.6	1038	6	Q28699	Q28699 oryctolagus
25	98.5	5.6	1935	4	Q9H1D5	Q9h1d5 homo sapien
26	98.5	5.6	2763	13	Q9YGV2	Q9ygv2 figu rubrip
27	98	5.6	561	1	Q9V2L7	Q9v2l7 pyrococcus
28	98	5.6	864	10	Q9T033	Q9t033 arabidopsis
29	98	5.6	980	14	P90478	P90478 feline sync
30	97.5	5.5	758	1	Q9UY62	Q9uy62 pyrococcus
31	97.5	5.5	982	14	Q56861	Q56861 feline foam
32	97.5	5.5	1348	5	Q9VAD1	Q9vad1 drosophila
33	96.5	5.5	533	1	Q9YCP2	Q9ycp2 aeropyrum p
34	96.5	5.5	574	10	Q9FY55	Q9fy55 arabidopsis
35	96.5	5.5	729	2	Q85239	Q85239 yersinia en
36	96.5	5.5	1062	2	P95422	P95422 pseudomonas
37	96.5	5.5	1062	2	Q910Y8	Q910y8 pseudomonas
38	96.5	5.5	1935	6	Q9GKR1	Q9gkr1 sus scrofa
39	96	5.4	390	2	Q9JZF1	Q9jzf1 neisseria m
40	95.5	5.4	1020	2	Q9KLN4	Q9kln4 vibrio chol
41	95.5	5.4	1209	5	Q93620	Q93620 caenorhabdi
42	95.5	5.4	2471	14	Q9YTK3	Q9ytk3 ateline her
43	95	5.4	2253	13	P70012	P70012 xenopus lae
44	94.5	5.4	769	4	Q9H4E9	Q9h4e9 homo sapien
45	94	5.3	1778	3	Q90TK4	Q90tk4 schizosacch

## ALIGNMENTS

RESULT 1  
Q9H2P8 ID Q9H2P8 PRELIMINARY; PRT; 350 AA.  
AC Q9H2P8;  
DT 01-MAR-2001 (Tremblrel. 16, Created)  
DT 01-MAR-2001 (Tremblrel. 16, Last sequence update)  
DT 01-MAR-2001 (Tremblrel. 16, Last annotation update)  
DE HT028.  
OS Homo sapiens (Human).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
OX NCBI\_TaxID=9606;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC TISSUE=HYPOPHALAMUS;  
RA Yang Y., Xu X., Gao G., Xiao H., Chen Z., Han Z.;  
RL Submitted (MAR-2000) to the EMBL/GenBank/DBJ databases.  
DR EMBL; AF248966; AAG44564.1; .  
SQ SEQUENCE 350 AA; 39036 MW; 840A989ACC4E8DE8 CRC64;

Query Match 99.8%; Score 1759; DB 4; Length 350;  
Best Local Similarity 99.7%; Pred. No. 5.8e-126;  
Matches 349; Conservative 1; Mismatches 0; Indels 0; Gaps 0;  
QY 1 MAVFVLLALVAGVLGNEFSILKSPGSVYFRNGNMPIPGERIPDVAAALSMGFVSKEDLSW 60  
DB 1 MAVFVLLALVAGVLGNEFSILKSPGSVYFRNGNMPIPGERIPDVAAALSMGFVSKEDLSW 60  
QY 61 PGLAVGNLFRPRATVYMWVGNKALPPGSVISYPLENAVPPFSLDSVANSIHLFSEE 120  
DB 61 PGLAVGNLFRPRATVYMWVGNKALPPGSVISYPLENAVPPFSLDSVANSIHLFSEE 120  
QY 121 TPVYLQALPSEERYVYVGNKANSVEDLSVTLQRLNRLFOQNSVLSLPLNSLRNNEVD 180  
DB 121 TPVYLQALPSEERYVYVGNKANSVEDLSVTLQRLNRLFOQNSVLSLPLNSLRNNEVD 180  
QY 181 LLFSELQVLDHDISSLLSRHKLAKDHSPLYSLEAGLDIGKRYGDESEQFRDASKIL 240  
DB 181 LLFSELQVLDHDISSLLSRHKLAKDHSPLYSLEAGLDIGKRYGDESEQFRDASKIL 240  
QY 241 VDLOKQFADDMYSLYGGNAVVELTVYKSFDTSLIKRTTILEAKQAKNPASPNLAYKN 300

```

Db 241 VVALQKFAADMYSLYGNVAVELTVKSFDTSLIRKTRTILEAKRAKNPASPYNLAYKN 300
Qy 301 FEYSVFNWNLWIMIALALAVIITSYINWNNWDPGYSIIYRMTNOKIRMD 350
Db 301 FEYSVFNWNLWIMIALALAVIITSYINWNNWDPGYSIIYRMTNOKIRMD 350

RESULT 2
ID Q9UG89 PRELIMINARY; PRT; 330 AA.
AC Q9UG89;
DT 01-MAY-2000 (TrEMBLrel. 13, Created)
DT 01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DT 01-MAY-2000 (TrEMBLrel. 13, Last annotation update)
DE HYPOTHETICAL 37.0 KDA PROTEIN (FRAGMENT).
GN DKF2P564O0582.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=BRAIN;
RA Pouska A., Klein M., Mewes H.W., Gassenhuber J., Wiemann S.;
RL Submitted (MAY-1999) to the EMBL/GenBank/DBJ databases.
DR EMBL; AL049929; CAB43210.1; -.
KW Hypothetical protein.
FT NON_TER
SQ SEQUENCE 330 AA; 37004 MW; 58098DCDF5BBCFE2 CRC64;

```

```

Query Match 94.6%; Score 1667; DB 4; Length 330;
Best Local Similarity 99.7%; Pred. No. 5.3e-119;
Matches 329; Conservative 1; Mismatches 0; Indels 0;

Qy 21 ILKSPGVFRNGNWPICGERIPDVAALSMGFSVKEDLSWPLAVGNLPHRPATVMVW 80
Db 1 ILKSPGVFRNGNWPICGERIPDVAALSMGFSVKEDLSWPLAVGNLPHRPATVMVW 60
Qy 81 KGVNKLALPGSVISYISPLENAVPSLDSVANSTHLSFSEETPVVQLAPSEERVYMGKA 140
Db 61 KGVNKLALPGSVISYISPLENAVPSLDSVANSTHLSFSEETPVVQLAPSEERVYMGKA 120
Qy 141 NSVFEDLSVTLRLRNRLFOENSVLSSLPNLSLRNNEVDLLFLSELQVLDHDISSLSRH 200
Db 121 NSVFEDLSVTLRLRNRLFOENSVLSSLPNLSLRNNEVDLLFLSELQVLDHDISSLSRH 180
Qy 201 KHLAKDHSPLISLELAGLDEIGKRYGEDSEQFRDASKILVDALQKPADDMYSLYGGNAV 260
Db 181 KHLAKDHSPLISLELAGLDEIGKRYGEDSEQFRDASKILVDALQKPADDMYSLYGGNAV 240
Qy 261 VELTVKSFDTSLIRKTRTILEAKRAKNPASPYNLAYKNFEYSVFNWNLWIMIALALA 320
Db 241 VELTVKSFDTSLIRKTRTILEAKRAKNPASPYNLAYKNFEYSVFNWNLWIMIALALA 300
Qy 321 VLIITSYINWNNWDPGYSIIYRMTNOKIRMD 350
Db 301 VLIITSYINWNNWDPGYSIIYRMTNOKIRMD 330

```

```

RESULT 3
ID Q9VHG4 PRELIMINARY; PRT; 320 AA.
AC Q9VHG4;
DT 01-MAY-2000 (TrEMBLrel. 13, Created)
DT 01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DT 01-MAY-2000 (TrEMBLrel. 13, Last annotation update)
DE CG8444 PROTEIN.
GN CG8444.
OS Drosophila melanogaster (Fruit fly).
OC Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
OC Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;

```

```

OC Ephydroidea; Drosophilidae; Drosophila.
OX NCBI_TaxID=7227;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=BERKELEY;
RX MEDLINE=20196006; PubMed=10731132;
RA Adams M.D., Celniker S.E., Holt R.A., Evans C.A., Gocayne J.D.,
RA Amanatides P.G., Scherer S.E., Li P.W., Hoskins R.A., Galle R.F.,
RA George R.A., Lewis S.E., Richards S., Ashburner M., Henderson S.N.,
RA Sutton G.G., Wortman J.R., Yandell M.D., Zhang Q., Chen L.X.,
RA Brandon R.C., Rogers Y.-H.C., Blazef R.G., Champe M., Pfeiffer B.D.,
RA Wan K.H., Doyle C., Baxter E.G., Heit G., Nelson C.R., Miklos G.L.G.,
RA Abell J.F., Aghayani A., An H.-J., Andrews-Pfannkoch C., Baldwin D.,
RA Ballew R.M., Basu A., Baxendale J., Bayraktaroglu L., Beasley E.M.,
RA Beeson K.Y., Benos P.V., Berman B.P., Bhandari D., Bolshakov S.,
RA Borkova D., Botchan M.R., Bouck J., Brokstein P., Brotter P.,
RA Burtis K.C., Busam D.A., Butler H., Cadieu E., Center A., Chandra I.,
RA Cherry J.M., Cawley S., Dahlke C., Davenport L.B., Davies P.,
RA de Pablo B., Delecher A., Deng Z., Mays A.D., Dew I., Dietz S.M.,
RA Dodson K.J., Evangelista C.C., Ferraz C., Ferreira S., Fleischmann W.,
RA Fostler C., Gabrielian A.E., Garg N.S., Gelbart W.M., Glasser K.,
RA Glodek A., Gong F., Gorrell J.H., Gu Z., Guan P., Harris M.,
RA Harris N.L., Harvey D., Heiman T.J., Hernandez J.R., Houck J.,
RA Hostin D., Houston K.A., Howland T.J., Wei M.-H., Ibegwam C.,
RA Jalali M., Kalush F., Karpen G.H., Ke Z., Kennison J.A., Ketchum K.A.,
RA Kimmel B.E., Kodira C.D., Kraft C., Kravitz S., Kulp D., Lai Z.,
RA Lasko P., Lei Y., Levitsky A.A., Li J., Li Z., Liang Y., Lin X.,
RA Liu X., Mattel B., McIntosh T.C., McLeod M.P., McPherson D.,
RA Merkulov G., Milshina N.V., Mobarry K., Morris J., Moshrefi A.,
RA Mount S.M., Moy M., Murphy B., Murphy L., Muzny D.M., Nelson D.L.,
RA Palazzolo M., Pittman G.S., Pan S., Pollard J., Puri V., Reese M.G.,
RA Reinert K., Remington K., Saunders R.D.C., Scheeler F., Shen H.,
RA Shue B.C., Siden-Kiamos I., Simpson M., Skupski M.P., Smith T.,
RA Spier E., Spradling A.C., Stapleton M., Strong R., Sun E.,
RA Svirskas R., Tector C., Turner R., Venter E., Wang A.H., Wang X.,
RA Wang Z.-Y., Wassarman D.A., Weinstock G.M., Weissbach J.,
RA Williams S.M., Woodage T., Worley K.C., Wu D., Yang S., Yao Q.A.,
RA Ye J., Yeh R.-F., Zaveri J.S., Zhan M., Zhang G., Zhao Q., Zheng L.,
RA Zheng X.H., Zhong F.N., Zhong W., Zhou X., Zhu S., Zhu X., Smith H.O.,
RA Gibbs R.A., Myers E.W., Rubin G.M., Venter J.C.;
RT "The genome sequence of Drosophila melanogaster.";
RL Science 287:2185-2195(2000).
DR EMBL; AE003682; AA054350.1; -.
DR FlyBase; FBgn0037671; CG8444.
SQ SEQUENCE 320 AA; 34421 MW; 401BCEAAA4FBA7E5 CRC64;

```

```

Query Match 18.5%; Score 327; DB 5; Length 320;
Best Local Similarity 26.3%; Pred. No. 4.9e-17;
Matches 94; Conservative 74; Mismatches 142; Indels 48; Gaps 11;

Qy 1 MAYVVLLALVAGV-LGNESILKSPGVFRNGNWPICGERIPDVAALSMGFSVKEDLS 59
Db 2 LRVFVIFSLFTAAINASGEFTVLRNPKAISPK-GNDALSHYVGVGVYASMGNAVSGDTN 60
Qy 60 WPGIAYGNLFRPRATVMVWVKVKNLALPPGVSISYISPLENAVPSLDSVANSTHLSFSE 119
Db 61 WNGLTINDPFLNLAGVILVHVQIGHVT-TAGNVKTYELTGS---GTASLNAL----- 110
Qy 120 ETPVVLQAPSEERVYMGKANSVFDLSVTILRLRNRLFOENSVLSSLPNLSLRNNEV 179
Db 111 -----AAAEAFANPEPCDIN-----FEQFDGQVQAKS-CFGDFEAPAAKPTKHLNPS--- 157
Qy 180 DLLFSELQVLDHDISSLSLRHKLADHSDPLYSLELAGLDEIGKRYGEDSEQFRDASKI 239
Db 158 --LHTADKQFLQEVGFINSAAHIAEMAKPSNVLMLRVSDGVGAHGEKSVAAVEANKL 215
Qy 240 LVDALOKPADDMYSLYCGNAVVELTVKSFDTSL-IRKTRTILEAKQAK-----NPASP 292
Db 216 LSAATSRLL-----AASQKSSDSVLFVQTTEKQVAAASRAKRTDTAAATTPN 261

```







```

Db 297 KVLGSMNNKDDITSIQANGKLSVOQFKLCICILAFQHPASLIRLEFDC----- 344
QY 232 QPRASKILVDALQFADDMYSLYGGNAVVELVTVKSFDTSLIRKTRTILEAKQAKNPAS 291
Db 345 ---SISRIEPKSLQKQVHTQVILSRNQITQINAVDDFAFSQL-PMLTSLDLSNRLES 400
QY 292 PYNLAY 297
Db 401 PSNVIY 406
RESULT 11
O25885
ID O25885 PRELIMINARY; PRT; 412 AA.
AC O25885;
DT 01-JAN-1998 (TRENBLrel. 05, Created)
DT 01-JAN-1998 (TRENBLrel. 05, Last sequence update)
DT 01-MAR-2001 (TRENBLrel. 16, Last annotation update)
DE HYPOTHETICAL 47.1 KDA PROTEIN.
GN HP1327.
OS Helicobacter pylori (Campylobacter pylori).
OC Bacteria; Proteobacteria; epsilon subdivision; Helicobacter group;
OC Helicobacter.
OX NCBI_TaxID=210;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=26695 / ATCC 700392;
RX MEDLINE=97394467; PubMed=9252185;
RA Tomb J.-F., White O., Kerlavage A.R., Clayton R.A., Sutton G.G.,
RA Fleischmann R.D., Ketchum K.A., Klenk H.-P., Gill S., Dougherty B.A.,
RA Loftus B., Richardson D., Zhou L., Kirkness E.F., Peterson S.,
RA McKenney K., Fitzgerald L.M., Lee N., Adams M.D., Hickey E.K.,
RA Berg D.E., Gocayne J.D., Utterback T.R., Peterson J.D., Kelley J.M.,
RA Cotton M.D., Weidman J.M., Fujii C., Bowman C., Watthey L., Wallin E.,
RA Hayes W.S., Borodovsky M., Karp P.D., Smith H.O., Fraser C.M.,
RA Venter J.C.;
RT "The complete genome sequence of the gastric pathogen Helicobacter
RT pylori."
RL Nature 388:539-547(1997).
DR EMBL; AE000634; AAC08377.1; -.
DR TIGR; HP1327; -.
DR InterPro; IPR003423; -.
DR Pfam; PF02321; OEP; 1.
KW Hypothetical protein.
SQ SEQUENCE 412 AA; 47115 MW; 1724D84A9BFA15DD CRC64;

```

```

Query Match 5.9%; Score 104; DB 2; Length 412;
Best Local Similarity 22.0%; Pred. No. 6.8;
Matches 70; Conservative 55; Mismatches 101; Indels 92; Gaps 15;
QY 48 LSMGFSVKEDLWPGVAGNLFH-RPRATVMVMGVNKLALPPGVSIVSYPLEN----- 100
Db 97 MSLGSGKVDLNGKKLTQSMIDLEKKILEKTKKQOLAI---SLMNGIENYKNOQE 153
QY 101 --AVPFSLOSANSIHSLSFETPVVLQAPSE-ERVVMGVKANSVFEDLSVTLRLNR 157
Db 154 IELLTKAIKNLTYQANSSSPNLIATIAKLEILKSLQLEIKKNLEALSASHYSGEL 213
QY 158 LFOENSVLSPNLSLRNNEVDLLFLSELQVLHDISSLSRHKHLAKDHPDLSLELA 217
Db 214 AFKENELLSIAPKN-FEFNRE-----QELHNISA-----TNYDIAIA 249
QY 218 GLDICKRGYGESEQFRDASK--ILVD-----ALQKFADDMYS-----LYGG 257
Db 250 RLDE-----EKSKDITLAKKSFLVDNVVTGVYFKSKQYVYNDMFSIALSIPPIYK 303
QY 258 NAVVELTVKSFDT-----SLIRKTTI---LEA-----KOAKNPA 290
Db 304 QA--KLVEOKKESILVFKSEVENTKNKTHLALKILKLETLQKNLESINKIKONEKTA 361
QY 291 SPYNLAYKNFEYSVFN 308

```

```

Db 362 QIVALDLKSGNDYNNAYN 379
RESULT 12
O68717
ID O68717 PRELIMINARY; PRT; 732 AA.
AC O68717;
DT 01-AUG-1998 (TRENBLrel. 07, Created)
DT 01-AUG-1998 (TRENBLrel. 07, Last sequence update)
DT 01-MAR-2001 (TRENBLrel. 16, Last annotation update)
DE PROTEIN KINASE A HOMOLOG.
GN YPKA.
OS Yersinia pestis.
OG Plasmid pCD1.
OC Bacteria; Proteobacteria; gamma subdivision; Enterobacteriaceae;
OC Yersinia.
OX NCBI_TaxID=632;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=KIM5;
RA Hu P., Elliott J., McCready P., Skowronski E., Garnes J.,
RA Kobayashi A., Carrano A.V., Brubaker R., Garcia E.;
RL Submitted (MAR-1998) to the EMBL/GenBank/DBJ databases.
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=KIM5;
RX MEDLINE=98427122; PubMed=9746557;
RA Perry R.D., Straley S.C., Fetherston J.D., Rose D.J., Gregor J.,
RA Blatner F.R.;
RT "DNA sequencing and analysis of the low-Ca2+-response plasmid pCD1 of
RT Yersinia pestis KIM5."
RL Infect. Immun. 66:4611-4623(1998).
CC -I- SIMILARITY: TO THE SER/THR FAMILY OF PROTEIN KINASES.
DR EMBL; AF053946; AAC62602.1; -.
DR EMBL; AF074612; AAC69765.1; -.
DR InterPro; IPR000719; -.
DR InterPro; IPR002290; -.
DR PRINTS; PR01373; YERSSTKINASE.
DR PROSITE; PS0011; PROTEIN_KINASE_DOM; 1.
DR PROSITE; PS00108; PROTEIN_KINASE_ST; 1.
DR SMART; SM00220; S_TKc; 1.
KW ATP-binding; Kinase; Plasmid; Serine/threonine-protein kinase;
KW Transferase.
SQ SEQUENCE 732 AA; 81732 MW; 1ECB0099EE2D3430 CRC64;

```

```

Query Match 5.9%; Score 103.5; DB 2; Length 732;
Best Local Similarity 21.2%; Pred. No. 18;
Matches 73; Conservative 51; Mismatches 103; Indels 117; Gaps 17;
QY 12 AGVLGNFESILKSGSVVFRNGNPIPIGERIPDVAALSM-----GFSVKEDLSWPG 62
Db 264 AGVVHNDI---KFGNVVFRAS---GE--PVYIDLGLHSRSGEQPKGFT--ESFKAPE 311
QY 63 LAVGNLFHRPRATVMVMV-----KGVNKLALPPGVSIVSYPLENAPV 103
Db 312 LGVNLGASEKSDVFLVVTLLHCIEGKPEIKPNQGLRFTITSEPAHVD---ENGYP 368
QY 104 FSLDSVANSIHSLSFETPVVQLA-----PSEERYVMGVKANSVFEDLSVTLRLNR 157
Db 369 IHRPGIA-GVETATVTRFITDILGVASDRSPDSNEARLHFLSDGTIDESAQI--LKDT 425
QY 158 LFOENSVLSPNLSLRNNEVDLLFLSELQVLHDISSLSRHKHLAKDHPDL----- 211
Db 426 LTGESMPSLT-----DVRITP-KKLRELSDDLRLHLSSAATKOLDMGVLS 472
QY 212 YSLELAGLDEIGKRYGSEDSQFRDASKILVDALQKFADDMYSLYGGNAVVELTVKSFDT 271
Db 473 LDTMLKVALDKAERGGVDKQ-----AKQAKNPASPNLAYKNFEYSV 498
QY 272 SLIRKTRTILE-----AKQAKNPASPNLAYKNFEYSV 306

```



Db 351 VDISNIFYGELTPILRRFR 370

## RESULT 15

Q9ZDN3 PRELIMINARY; PRT; 470 AA.  
AC Q9ZDN3;  
DT 01-MAY-1999 (TrEMBLrel. 10, Created)  
DT 01-MAY-1999 (TrEMBLrel. 10, Last sequence update)  
DT 01-MAY-2000 (TrEMBLrel. 13, Last annotation update)  
DE GUANOSINE PENTAPHOSPHATE PHOSPHOHYDROLASE (GPPA).  
GN RP294.  
OS Rickettsia prowazekii.  
OC Bacteria; Proteobacteria; alpha subdivision; Rickettsiales;  
OC Rickettsiaceae; Rickettsieae; Rickettsia.  
OX NCBI\_TaxID=782;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN=MADRID E;  
RX MEDLINE=99039499; PubMed=9823893;  
RA Andersson S.G.E., Zomorodipour A., Andersson J.O.,  
RA Sacheritz-Ponten T., Alsmark U.C.M., Podowski R.M., Naeslund A.K.,  
RA Eriksson A.-S., Winkler H.H., Kurland C.G.;  
RT "The genome sequence of Rickettsia prowazekii and the origin of  
RT mitochondria."  
RL Nature 396:133-140(1998).  
DR EMBL; AJ235271; CAA14755.1; .  
SQ SEQUENCE 470 AA; 53780 MW; 74F7E97599D02903 CRC64;

## Query Match

Best Local Similarity 5.8%; Score 103; DB 2; Length 470;

Matches 65; Conservative 56; Mismatches 123; Indels 56; Gaps 12;

Qy 45 VAALSMGFSYKEDLSWPGIAGVGNLFHRPRATVMVMKGVNKL-ALPPGVSISYPLENAV 103  
Db 117 ISGISDAFGIVADLGGSLTAHYN-----KVGKLSLPLGTKI----- 157  
Qy 104 FSLDSVANS-----IHSLSFSETPVVLQAPSEERVYVMGKA-----NSVFEDLSVTL 151  
Db 158 -----IANSFSDVGLITKMLEEFGVA-----HYPNLYLIGGALRLMSRIYMESINYPL 207  
Qy 152 ROLRNLFOENSYSLSPLNSLRNNEVDLLFLSELOVLHDISSLSRKHKLAKDHPDL 211  
Db 208 KNLHN--FEINRVFEFLYLEKLSQIDKLSYEQKAINYNAVLI---KAMIKVFESEK 262  
Qy 212 YSLELAGLDEIGRYGEDSEQFRASKIL-----VDALQKFADDMYSLYGNNAVVELTVK 267  
Db 263 IIISNYGLKE-GVRF--DSLPHYETEKDIIYERVKLVNFDNRNICKIERKIEALQYLLIN 319  
Qy 268 SFDTSLRKTRTILEAKQAKN--PASPYNLAYKNPEKSVFNMVLMIALALAVITTS 325  
Db 320 SDATLLIIELAIQLAQNKNIDKTLRANFVSEFILLSDDIPFSHRQRLMLGIALVTYTA 379

Search completed: July 3, 2001, 16:20:24  
Job time: 156 sec